The Impact of Social Norms on Cohesion and (De)Polarization

Edited by Miranda Lubbers, Marcin Bukowski, Oliver Christ, Eva Jaspers, and Maarten van Zalk

cogitatio

Table of Contents

Table of Contents

EDITORIAL

The Impact of Social Norms on Cohesion and (De)Polarization

Miranda J. Lubbers [1] 🄳, **Marcin Bukowski** [2] 🄳, **Oliver Christ** [3] 🄳, **Eva Jaspers** [4] 🄳, and **Maarten van Zalk** [5] 🄳

[1] Department of Social and Cultural Anthropology, Universitat Autònoma de Barcelona, Spain
[2] Institute of Psychology, Jagiellonian University, Poland
[3] Faculty of Psychology, FernUniversität in Hagen, Germany
[4] Department of Sociology, Utrecht University, The Netherlands
[5] Department of Developmental Psychology, University of Osnabrück, Germany

Correspondence: Miranda J. Lubbers (mirandajessica.lubbers@uab.cat)

Abstract

In recent years, political and social polarization has increased across many societies, evolving from mere issue-based disagreements into affective polarization, in which citizens dislike and distrust members of opposing groups. This trend undermines social cohesion and the effective functioning of democratic institutions. Despite extensive interdisciplinary research into polarization, the role of social norms—shared expectations about typical and appropriate behavior—in mitigating such divisions remains underexamined. This thematic issue seeks to address this gap by investigating how social norms shape intergroup dynamics in polarized contexts. To frame the contributions, this introductory article first outlines the concepts of polarization and social norms and then briefly reviews the literature on the role of norms in polarization and depolarization. Subsequently, we introduce the contributions included in this thematic issue, which explore four central themes: (a) the role of social norms in fostering tolerance and depolarization; (b) the association between norm deviations and non-normative behaviors and political polarization; (c) the negotiation versus contestation of social norms by competing groups; and (d) the influence of social networks on intergroup attitudes and behaviors that can facilitate depolarization processes. We conclude with reflections on future research directions.

Keywords

intergroup relationships; interventions; polarization; social cohesion; social norms

---

1. Introduction

Over the past few decades, societies around the world have experienced rising political polarization (McCoy et al., 2018; Phillips, 2022; Reiljan, 2019). Ideological polarization, typically defined as the growing distance between political parties or candidates in terms of policy positions and values (Harteveld et al., 2022), can, to some extent, benefit democracy by clarifying party choices, stimulating democratic debate, and encouraging civic engagement. However, it turns harmful when "the normal multiplicity of differences in a society increasingly align along a single dimension," leading people to view politics and society through a stark "us versus them" lens (McCoy et al., 2018, p. 16; cf. Jost et al., 2022; McCoy & Somer, 2019).

Three interrelated tendencies make this rise in polarization particularly concerning. First, political identities are increasingly aligned with other salient social identities, such as ethnicity, migration status, gender, sexual orientation, and social class. For example, Iyengar et al. (2019) observed that "White evangelicals...are overwhelmingly Republican today, and African Americans overwhelmingly identify as Democrats" (p. 134), illustrating the correlation between political affiliation and race and religion in the US. This alignment extends into lifestyle domains, where preferences for leisure activities and consumption choices are increasingly politicized, a phenomenon popularly known as the "culture wars" (Hunter, 1991; cf. DellaPosta et al., 2015). Lifestyle politics, defined as the "politicization of everyday life choices, including ethically, morally or politically inspired decisions about, for example, consumption, transportation, or modes of living" (de Moor, 2017, p. 181), reinforces this dynamic. As a result, a "stereotypical world" emerges, a world of 'latte liberals' and 'bird-hunting conservatives'" (DellaPosta et al., 2015, p. 1475). In such a world, politically opposed groups have less common ground.

Second, and relatedly (Iyengar et al., 2012, 2019), polarization today is not limited to ideological disagreement but increasingly takes the form of *affective* polarization, defined as the tendency to dislike, distrust, and avoid those with opposing ideologies (Gidron et al., 2023; Iyengar et al., 2019). Affective polarization has intensified over time (Gidron et al., 2020; Iyengar et al., 2019) and is characterized by emotional hostility and social distancing between partisans (Balcells & Kuo, 2022; Hobolt et al., 2021). It is particularly pronounced in societies with high income inequality and unemployment (Gidron et al., 2020), where structural grievances may amplify intergroup resentment.

Third, political and media actors, including populist parties, often deliberately exploit affective polarization (Davis et al., 2024) to mobilize their base and consolidate support, thereby exacerbating divisions. Populist parties typically use discursive frames that distinguish between the people they claim to protect ("us") and a corrupt elite they claim to fight ("them"; Mudde & Rovira Kaltwasser, 2017). Scholars have therefore argued that such parties inherently foster affective polarization by positioning themselves as the sole defenders of democracy while portraying their rivals as threats (e.g., Roberts, 2021; Urbinati, 2019). By doing so, they justify the restriction of their rivals' civil liberties and compromise fair elections. In addition, populists tend to redefine who belongs to the democratic community, often excluding immigrants, other minorities, and political opponents from being seen as legitimate citizens. As a result, tolerance declines, negative stereotypes about political opponents are reinforced, and social interactions outside the political arena are discouraged. In addition, populist rhetoric often provokes a reaction from opponents, pushing them to adopt similarly divisive tactics (Stavrakakis, 2018), such as portraying populist voters as ignorant or claiming exclusive ownership of democratic values. Davis et al. (2024) show that affective polarization is higher in countries with a strong presence of populism, not only among its supporters but also among its critics.

Together, these tendencies illustrate how polarization has spilled over from the political realm into everyday social life, shaping how individuals perceive others and interact with them across categorical boundaries of class, gender, sexual orientation, ethnicity, and religion (Castle, 2018). These divisions erode mutual understanding, reduce tolerance, and hinder intergroup contact and cooperation. At the societal level, they threaten the functioning of democracies (McCoy et al., 2018) by fueling hate speech and political violence (Suarez Estrada et al., 2022), impeding collective problem-solving, and contributing to democratic backsliding (Orhan, 2021).

Despite extensive research on the causes and consequences of polarization (Iyengar et al., 2019) and growing interest in interventions aimed at reducing affective polarization (Huddy & Yair, 2021; Levendusky, 2018; Tausch et al., 2024; Voelkel et al., 2022), relatively little attention has been paid to the role of social norms. Yet, social norms—shared expectations about common and appropriate behavior—are powerful regulators of intergroup dynamics. They can shape how individuals express disagreement, regulate negative sentiment toward other social groups (Iyengar & Westwood, 2014; Meleady, 2021), and stimulate the willingness to engage respectfully with people one disagrees with. As such, norms may either mitigate or exacerbate the harmful effects of polarization. However, we still know relatively little about how they are perceived, transmitted, contested, or enforced in polarized contexts, and under what conditions they can help mitigate the harmful effects of polarization.

This thematic issue addresses this gap by investigating how the transmission of and conformity to prosocial norms can promote tolerance and foster the willingness to interact and collaborate with other social groups in polarized contexts, and if and how norms can decrease polarization. It was organized within the framework of the research project Inclusivity Norms to Counter Polarization in European Societies (INCLUSIVITY), an international collaborative project funded by the Volkswagen Foundation that involved the guest editors of this thematic issue. This project investigates inclusivity norms, defined as social group norms that promote equality-based respect, dialogue, and unity. It aims to identify what the core ingredients of such norms should be to prevent the corrosive consequences of polarization on the social fabric of European societies (e.g., Ciordia et al., in press; Richters et al., 2025; Schäfer et al., 2024; Shani et al., 2023; Simons et al., 2025). This thematic issue presents results from the INCLUSIVITY project as well as contributions from other scholars working on similar topics from psychology, sociology, political sciences, civilization studies, and anthropology. By exploring the impact of norms on (de-)polarization, this thematic issue aims to gain new insights into creating a more cohesive and inclusive society.

In the following two sections, we introduce the concept of social norms and their key dimensions, and then review the literature on their role in polarization and depolarization. Subsequently, we present the contributions included in this thematic issue. We conclude with a forward-looking reflection on possible directions for future research.

2. Social Norms

Social norms are generally understood as the informal rules that govern individual behavior within groups and societies (Bicchieri et al., 2023; Coleman, 1990; Hecter & Opp, 2001). They act as a "grammar of social interactions" (Bicchieri, 2006; Bicchieri et al., 2023) or an unwritten "code of conduct" (Geber & Sedlander, 2022), informing group members about what is considered acceptable and unacceptable behavior in specific

situations. Thus, by prescribing and proscribing particular behaviors in a wide range of social situations, from how and when to respond to disagreement to when to help others, norms facilitate smooth social interaction and contribute to group cohesion.

For individuals to act upon group norms, they first need to perceive them. Geber and Sedlander (2022) call this macro-micro link from collective norms to individual norm perceptions a "communication link" (p. 4), where individuals infer group norms from the verbal and non-verbal cues provided by referents. A reference group is defined as "the relevant others whose behavior and (dis)approval matter in sustaining the norm" (Legros & Cislaghi, 2020, p. 75), which includes mostly other group members, but also authorities, institutions, and media content. This process begins early in life through socialization by parents, teachers, and peers, and continues throughout adulthood, as individuals constantly update their impressions of what behaviors are typical and valued within their social networks, wider social groups, and society at large (Tankard & Paluck, 2016). Given this ongoing process of social learning and updating, both social norms and norm perceptions remain subject to change. Scholars often distinguish between individuals' perceptions of "descriptive norms" (common behaviors of the reference group in a given situation) and "injunctive norms" (behaviors that are valued or approved by the reference group for a given situation; Cialdini et al., 1990).

Social norms are a key driver of human behavior. According to social identity theory (Tajfel & Turner, 1979), individuals derive part of their self-concept from group memberships and therefore, they tend to be motivated to internalize group norms and abide by them (cf. González, 2025); norms even become part of one's identity (Abrams & Hogg, 1990; Meleady, 2021). Social norm interventions are therefore effective in stimulating desirable outcomes in many areas: simply evoking a social norm enhances the probability that people reuse towels in hotels, eat healthier, vote in general elections, give to charity, and care for the environment (research summarized by Meleady, 2021). Evoking social norms also encourages intergroup contact: Studies have shown that individuals are more likely to engage in intergroup contact when they perceive it as socially approved or expected, highlighting the importance of descriptive and injunctive norms in guiding individual but also collective behavior (Prentice & Paluck, 2020; Tankard & Paluck, 2016). However, norm violations are also frequent. Norm compliance can also be enforced by the group or society exerting normative pressure and sanctions, including gossip, ostracism, group exclusion, and punishment (Eriksson et al., 2021).

3. Social Norms, Affective Polarization, and Depolarization

Despite the growing literature on affective polarization, the role of social norms remains relatively underexplored (Balafoutas et al., 2024; Cole et al., 2025). Most research has focused on psychological, institutional, or media-related drivers of polarization, including partisan identity, misinformation, and echo chambers (Huddy & Yair, 2021; Iyengar et al., 2019). Interventions to reduce affective polarization have often emphasized empathy-building, perspective-taking, or exposure to counter-attitudinal views (Voelkel et al., 2022). However, these approaches pay relatively little attention to social norms.

The limited research that does exist on the role of social norms in affective (de-)polarization suggests that they can both mitigate and exacerbate polarization, depending, firstly, on how the members of a group perceive ingroup norms regarding interactions with outgroups (González, 2025). On the one hand, social norms may reinforce affective polarization if the ingroup is perceived as avoiding contact with outgroup

members (descriptive norm) or disapproving of such contacts (injunctive norm). On the other hand, social norms may reduce affective polarization if the ingroup is perceived as valuing open and respectful dialogue and tolerance, even toward outgroup members (Iyengar & Westwood, 2014; Meleady, 2021).

Recent work highlights the importance of perceived ingroup norms in shaping the success of depolarization interventions (You & Lee, 2024). In polarized settings, individuals often significantly overestimate their ingroup's negative sentiments toward the outgroup, and correcting these misperceptions has been shown to reduce polarization (You & Lee, 2024). In addition, evoking a more overarching norm of open-mindedness has been shown to diversify engagement with news content (Wojcieszak et al., 2020). Normative appeals to a broadly shared social identity—such as priming shared national identity or emphasizing common goals—have also been found to reduce partisan animosity in experimental settings (Huddy & Yair, 2021; Levendusky, 2018). Similarly, priming external threats reduced outgroup discrimination in experiments (Kaba et al., 2024).

Yet, recent research also suggests other factors that may reduce the effectiveness of norm interventions in highly polarized settings. For instance, when groups are affectively highly polarized, knowledge about outgroup norms may provoke reactance rather than alignment. For instance, experimental evidence from the Covid-19 pandemic shows that stressing the low mask usage among Trump supporters did not leave Biden supporters unaffected, but rather increased their own intentions to wear masks (Rand & Yoeli, 2024). Furthermore, as indicated in Section 1, not only do norms affect polarization, severe polarization can also harm democratic norms such as the norm of political tolerance—the belief that all citizens, including outgroup members, deserve equal rights such as rights to free speech, protest, and voting (Kingzette et al., 2021). First, partisans can selectively endorse constitutional protections depending on which party is in power. Second, when the outgroup is seen as a threat, norms of political tolerance lose support. Finally, in highly polarized contexts, norms can be rapidly redefined in response to perceived threats, as illustrated by the surge in Spanish nationalist displays following the Catalan independence referendum. This shift was not driven by changing preferences but by a transformation in perceived norms, which made previously stigmatized behaviors more socially acceptable (Dinas et al., 2024). Together, these findings indicate that social norms may lose their unifying function, become politicized, and even reinforce division in highly polarized contexts. Therefore, future research and interventions must account for the dynamics of norms in such settings.

Not only does the effectiveness of norms vary with the level of polarization and perceived threat, the broader intergroup context also matters; for instance, the severity and nature of intergroup conflict—such as whether two groups must co-exist in shared spaces, such as Hutus and Tutsis, or co-exist in neighboring spatial contexts with no need for group interaction, or can remain spatially separate to a large extent, such as Israelis and Palestinians (Ron et al., 2017). Furthermore, the relative size of in- and outgroups can affect norms toward positive intergroup contact (Christ et al., 2014).

Despite these advances, much remains unknown about how social norms are formed, transmitted, and contested in real-world polarized environments. This is particularly true in digital spaces, where norms may evolve rapidly and be enforced by informal or algorithmic mechanisms. Cross-cultural variation in norm sensitivity and enforcement also remains underexplored, limiting the generalizability of existing findings. Addressing these gaps is crucial for developing more context-sensitive and norm-aware approaches to depolarization.

4. Contributions to the Thematic issue

This thematic issue advances research on the role of social norms in (de-)polarization in four regards. First, several contributions study the role of social norms in fostering tolerance and depolarization. Estevan-Reina et al. (2025) tested whether the association between personal respect norms and tolerance of opposing opinions is moderated by the extremity of one's own opinion, the strength of disapproval of the opposing opinion, and the perceived threat from the out-group. Their results, based on a large-scale, cross-European survey, reveal that the association between personal respect norms and tolerance is unaffected by these three moderators. Furthermore, the pattern of results largely replicates across twelve countries and six controversial social topics, even when considering differences in political views. Thus, this article provides solid evidence for the robustness of the link between equality-based respect norms and tolerance for opposing opinions.

Gurbisz et al. (2025) examine the association between equality-based respect norms and two levels of tolerance, namely general tolerance, understood as an abstract principle, and specific tolerance, as applied in concrete, real-world situations. They conducted three studies in Poland focusing on tolerance toward Jewish people. The results show that perceived prescriptive equality-based respect norms were positively linked to tolerant attitudes. Additionally, the findings reveal a possible overestimation of tolerant attitudes when assessed at an abstract level, compared to when measured through concrete examples, highlighting the importance of studying how norms translate into behavior in specific contexts.

Ackermann et al. (2025) investigated different types of social norms that promote civic volunteering, a key indicator of social cohesion. They specifically examined the differential effects of perceptions of participatory norms (e.g., citizens should be active in social or political organizations, and base their consumption choices on ethical, environmental, or political reasoning) versus solidarity norms (e.g., citizens should help others in need inside and outside their own country) on volunteering. They used a large population survey in Germany to test these associations. Their findings stress the importance of fostering both participatory and solidarity norms to stimulate volunteering, which in turn strengthens social cohesion in polarized contexts.

Nijs (2025) shows that informing individuals about the actual descriptive norm can help mitigate negative consequences of polarization, especially when these individuals misperceive social norms. In a pre-registered survey experiment conducted in the Netherlands, he tested whether providing participants with accurate information about the descriptive norm (i.e., actual differences in immigration attitudes within society) would reduce the association between perceptions of polarization and polarization worry. He found that this is indeed the case, but only among participants who correctly understood the provided information and recognized that their beliefs were inaccurate.

Finally, Jaspers et al. (2025) examined the role of descriptive norms encouraging interethnic dating on interethnic dating in adolescence. Using Dutch longitudinal survey data, they found that descriptive norms moderated the association between attitudes toward other ethnic groups and interethnic dating, such that adolescents with positive outgroup attitudes in wave 1 were more likely to engage in interethnic dating in waves 2 and 3 when descriptive norms were more supportive of it.

Second, several contributions to this thematic issue examine the association between, on the one hand, norm deviations and behaviors that do not conform to societal norms (i.e., non-normative behaviors), and on the other hand, political polarization. Lutterbach and Beelmann (2025) examined the psychological foundations of political polarization among German youth. They showed that norm deviation and intolerance shape polarized political attitudes and behavioral intentions (including, among other factors, the rejection of democracy and human rights, and an affinity for political violence), with distinct patterns observed across different demographic groups. These results highlight the important role of tailored educational and preventive interventions to reduce polarization risks and promote social cohesion among adolescents.

Kesberg et al. (2025) tested the hypothesis that perceived societal polarization reduces trust in the government, particularly when the government is perceived as posing a threat to the ingroup, and that low trust, in turn, increases intentions to engage in non-normative collective action. The authors experimentally manipulated polarization and threat in the context of a fictitious country, and then measured participants' intentions to engage in normative (peaceful), non-normative (vandalizing and disruptive), and extreme non-normative (violent or radicalized) collective action, as well as their trust in the government of that country. They found that perceived threat, but not polarization itself, significantly increased collective action intentions, especially non-normative ones. These results reveal the importance of factors beyond norms, such as perceived threat, in shaping collective behavior in polarizing settings.

Third, social norms are contested and negotiated by competing groups, particularly in times of social upheaval or political realignment. In polarized societies, competing groups may promote conflicting normative expectations, leading to norm fragmentation or even norm conflict. For instance, what one group views as "free speech" may be perceived by another as "hate speech." Understanding how norms are enforced, resisted, or redefined in such contexts is crucial for grasping their role in either exacerbating or mitigating polarization. In this realm, Kłodkowski et al. (2025) propose a conceptual framework for identity politics, in which dialogue and negotiation processes about micro-identities defined as "a deeply subjective sense of belonging to an exceptionally specific and narrowly defined group" (p. 5) play a prominent role. These processes of identity negotiation can shape norms and thereby either exacerbate or bridge political cleavages, separating or bringing together oppositional opinion groups on highly polarized topics. The authors use the case study of the national discourse in Poland about the TERF label (Trans-Exclusionary Radical Feminist) to illustrate a polarizing micro-identity strategy within feminism, and Polish politician Donald Tusk's electoral discourse to illustrate an integrative micro-identity strategy.

Bayarri-Toscano and Fernández-Villanueva (2025) explored how the circulation of memes on social media has contributed to the rise of far-right political parties in South American countries. Drawing on visual social media data and ethnographic fieldwork on recent electoral campaigns in Argentina, Brazil, and El Salvador, they explored the role of memes on social media and contextualized them in the affective climates surrounding electoral campaigns. Their analysis reveals that by transforming violent discourse into humorous memes, extreme right-wing parties normalized and legitimized violence against political adversaries, feminism, racialized persons, and people experiencing poverty, while deflecting accountability. In this context, humor functions as a mechanism for shaping social norms: it signals which forms of discourse are socially acceptable and influences public attitudes toward political violence. In doing so, memes become tools for reinforcing exclusionary norms and redefining the boundaries of permissible political expression.

Fourth, two articles in this thematic issue explore the impact of social networks, particularly their heterogeneity and the types of norms perceived within them (inclusive vs. exclusionary), on intergroup attitudes and behaviors that can facilitate depolarization (cf. Shepherd, 2017). Górska et al. (2025) focus on the role of diverse social networks in fostering tolerance toward opposing viewpoints. In a survey conducted in Poland, a highly polarized socio-political context, they examined whether network heterogeneity can be considered a distinct source of depolarization, separate from mere intergroup contact. Górska and colleagues show that having a more heterogeneous social network in terms of partisanship is indirectly associated with more positive attitudes toward political opponents. Additionally, this effect is independent of intergroup contact.

While it is well established that schools, families, peer groups, and digital platforms all play a role in shaping young people's early understanding of respectful or hostile political engagement, few studies have examined how early exposure to inclusive or exclusionary norms affects the development of political attitudes and intergroup behavior over time. Addressing this gap in a study in German secondary schools, Shani et al. (2025) innovatively employed a longitudinal approach to show that, whereas political ideological orientations fluctuate considerably during adolescence, youths' perceptions of social norms and their levels of affective polarization remain stable already from early adolescence onward. These findings underline the developmental fluidity and stability during adolescence, offering both opportunities and challenges for interventions. They suggest that effective depolarization strategies must account for the evolving nature of adolescent political socialization and the varying influence of school-based normative contexts.

5. Conclusions

The multidisciplinary contributions in this thematic issue demonstrate, both empirically and conceptually, how social norms can shape (de-)polarization. Drawing on evidence from diverse countries and topic areas, the studies highlight the central role of social norms in regulating intergroup attitudes, behaviors, and perceptions in polarized societies. One study shows that group-based differences in norm perceptions and polarization tendencies are already evident in early adolescence, even when political orientations remain fluid. Others highlight the robustness of equality-based respect norms in fostering tolerance, and the capacity of social norms to promote bridging behaviors, such as civic volunteering and interethnic dating. Conversely, the studies also draw attention to norm deviation as predictors of polarized behaviors. The contributions also emphasize the potential of norm-based interventions to reduce polarization and polarization anxieties, such as correcting misperceived descriptive norms. Furthermore, they reveal how contextual factors, such as perceived threat, social network heterogeneity, and identity politics in public discourse, moderate the influence of norms. Methodologically, the issue stressed the need to distinguish between abstract norm adherence and norm adherence in real-world contexts.

By bringing together interdisciplinary perspectives and diverse methodologies, this thematic issue paves the way for a broader research agenda on the role of social norms in fostering democratic behaviors and reducing polarization. We identify four key directions for future research. First, future work could consider the multiple environments in which prosocial norms emerge and are adopted and norm deviation discouraged, including schools and digital environments where norm violations can be amplified and rapidly disseminated. There is insufficient knowledge regarding how norms across different contexts interact. Ideally, such research would also contribute to a more multilevel understanding of social norms (Geber & Sedlander, 2022)—how they

operate across individual, group, institutional, and societal levels. Social network analysis could be further integrated in these initiatives to reveal how norms spread through social networks and how network structure influences the success of depolarization efforts.

Second, comparative cross-national studies are crucial for identifying the political, cultural, and institutional conditions that facilitate or hinder the acceptance of norms. While the contributions of this thematic issue offer valuable evidence, they primarily focus on so-called WEIRD (Western, educated, industrialized, rich, and democratic) societies. Extending empirical research to non-WEIRD contexts will provide a more complete understanding of cross-cultural variation in norm formation and enforcement and its intersection with polarization. Such an extension is also crucial for assessing the generalizability of existing findings across diverse sociopolitical environments.

Third, future research should investigate the mechanisms of norm perception, transmission, negotiation, and contestation in severely polarized environments as dynamic and context-dependent processes. Researchers could examine how norms are negotiated, reinterpreted, or resisted among competing groups, shaped by identity politics, and transformed in response to perceived threats or shifting social networks—processes that are central to understanding group alignment and resistance in polarized contexts. Both qualitative and quantitative research could contribute to such understanding. Furthermore, longitudinal studies could provide valuable insights into the lives of inclusive norms, or how they emerge, strengthen, stabilize, mutate, erode, and disappear, particularly in response to shifting political climates.

Finally, while norm-based interventions are promising for reducing polarization, their long-term efficacy remains uncertain. Most studies are short-term and experimental, leaving open questions about the durability of their effects. More intervention-based research is needed that tests the effectiveness of norm-based strategies for depolarization in real-world settings, both in the short and long term. By addressing these questions, future research can help build a more comprehensive and globally relevant framework for understanding how social norms can be leveraged to promote inclusive and resilient societies.

Conflict of Interests

The authors declare no conflict of interests.

References

Abrams, D., & Hogg, M. A. (Eds.). (1990). *Social identity theory: Constructive and critical advances*. Springer.

Ackermann, K., Mylius, J., Haussmann, A., & Wiloth, S. (2025). Living up to your own standards? Patterns of civic norms and volunteering in Germany. *Social Inclusion*, *13*, Article 10007. https://doi.org/10.17645/si.10007

Balafoutas, L., Dimant, E., Gächter, S., & Krupka, E. (2024). Social norms: Enforcement, breakdown & polarization. *European Economic Review*, *170*, Article 104885. https://doi.org/10.1016/j.euroecorev.2024.104885

Balcells, L., & Kuo, A. (2022). Secessionist conflict and affective polarization: Evidence from Catalonia. *Journal of Peace Research*, *60*(4), 604–618. https://doi.org/10.1177/00223433221088112

Bayarri-Toscano, G., & Fernández-Villanueva, C. (2025). "Funny weapons": The norms of humour in the construction of far-right political polarisation. *Social Inclusion*, *13*, Article 10211. https://doi.org/10.17645/si.10211

Bicchieri, C. (2006). *The grammar of society: The nature and dynamics of social norms*. Cambridge University Press.

Bicchieri, C., Muldoon, R., & Sontuoso, A. (2023). Social norms. In E. N. Zalta & U. Nodelman (Eds.), *The Stanford encyclopedia of philosophy*. Stanford University. https://plato.stanford.edu/archives/win2023/entries/social-norms

Castle, J. (2018). New fronts in the culture wars? Religion, partisanship, and polarization on religious liberty and transgender rights in the United States. *American Politics Research*, *47*(3), 650–679. https://doi.org/10.1177/1532673X18818169

Christ, O., Schmid, K., Lolliot, S., Swart, H., Stolle, D., Tausch, N., Al Ramiah, A., Wagner, U., Vertovec, S., & Hewstone, M. (2014). Contextual effect of positive intergroup contact on outgroup prejudice. *Proceedings of the National Academy of Sciences of the United States of America*, *111*, 3996–4000. http://doi.org/10.1073/pnas.1320901111

Cialdini, R. B., Reno, R. R., & Kallgren, C. A. (1990). A focus theory of normative conduct: Recycling the concept of norms to reduce littering in public places. *Journal of Personality and Social Psychology*, *58*, 1015–1026. http://doi.org/10.1037/0022–3514.58.6.1015

Ciordia, A., Targarona Rifa, N., & Lubbers, M. J. (in press). The relational toll of political involvement in polarized times: relationship decay within activists' personal networks. *American Behavioral Scientist*.

Cole, J. C., Gillis, A. J., van der Linden, S., Cohen, M. A., & Vandenbergh, M. P. (2025). Social psychological perspectives on political polarization: Insights and implications for climate change. *Perspectives on Psychological Science*, *20*(1), 115–141. https://doi.org/10.1177/17456916231186409

Coleman, J. S. (1990). *Foundations of social theory*. Harvard University Press.

Davis, B., Goodliffe, J., & Hawkins, K. (2024). The two-way effects of populism on affective polarization. *Comparative Political Studies*, *58*(1), 122–154. https://doi.org/10.1177/00104140241237453

de Moor, J. (2017). Lifestyle politics and the concept of political participation. *Acta Politica*, *52*(2), 179–197. https://doi.org/10.1057/ap.2015.27

DellaPosta, D., Shi, Y., & Macy, M. (2015). Why do liberals drink lattes? *American Journal of Sociology*, *120*(5), 1473–1511. https://doi.org/10.1086/681254

Dinas, E., Martínez, S., & Valentim, V. (2024). Social norm change, political symbols, and expression of stigmatized preferences. *The Journal of Politics*, *86*(2), 488–506. https://doi.org/10.1086/726951

Eriksson, K., Strimling, P., Gelfand, M., Wu, J., Abernathy, J., Akotia, C. S., Aldashev, A., Andersson, P. A., Andrighetto, G., Anum, A., Arikan, G., Aycan, Z., Bagherian, F., Barrera, D., Basnight-Brown, D., Batkeyev, B.,

Belaus, A., Berezina, E., Björnstjerna, M., . . . Van Lange, P. A. M. (2021). Perceptions of the appropriate response to norm violation in 57 societies. *Nature Communications*, *12*(1), Article 1481. https://doi.org/10.1038/s41467-021-21602-9

Estevan-Reina, L., Schäfer, L. F., Middendorf, W., Bukowski, M., van Zalk, M., & Christ, O. (2025). Testing the robustness of the association between personal respect norms and tolerance in polarized contexts. *Social Inclusion*, *13*, Article 10035. https://doi.org/10.17645/si.10035

Geber, S., & Sedlander, E. (2022). Communication as the crucial link: Toward a multilevel approach to normative social influence. *Studies in Communication Sciences*, *22*(2), 1–18. https://doi.org/10.24434/j.scoms.2022.02.005

Gidron, N., Adams, J., & Horne, W. (2020). *American affective polarization in comparative perspective*. Cambridge University Press.

Gidron, N., Adams, J., & Horne, W. (2023). Who dislikes whom? Affective polarization between pairs of parties in Western democracies. *British Journal of Political Science*, *53*(3), 997–1015. https://doi.org/10.1017/S0007123422000394

González, R. (2025). The role of norms as promoters of social change. *Political Psychology*, *46*(2), 455–481. https://doi.org/10.1111/pops.13025

Górska, P., Bulska, D., & Górski, M. (2025). The more, the merrier...: The effect of social network heterogeneity on attitudes toward political opponents. *Social Inclusion*, *13*, Article 10282. https://doi.org/10.17645/si.10282

Gurbisz, D., Potoczek, A., Bukowski, M., Estevan-Reina, L., & Christ, O. (2025). Mind the gap! Linking equality-based respect norms with general and specific tolerance. *Social Inclusion*, *13*, Article 10143. https://doi.org/10.17645/si.10143

Harteveld, E., Mendoza, P., & Rooduijn, M. (2022). Affective polarization and the populist radical right: Creating the hating? *Government and Opposition*, *57*(4), 703–727. https://doi.org/10.1017/gov.2021.31

Hecter, M., & Opp, K.-D. (Eds.). (2001). *Social norms*. Russel Sage Foundation.

Hobolt, S. B., Leeper, T. J., & Tilley, J. (2021). Divided by the vote: Affective polarization in the wake of the Brexit referendum. *British Journal of Political Science*, *51*, 1476–1493. https://doi.org/10.1017/S0007123420000125

Huddy, L., & Yair, O. (2021). Reducing affective polarization: Warm group relations or policy compromise? *Political Psychology*, *42*, 291–309. https://doi.org/10.1111/pops.12699

Hunter, J. D. (1991). *Culture wars: The struggle to define America*. Basic Books.

Iyengar, S., Lelkes, Y., Levendusky, M., Malhotra, N., & Westwood, S. J. (2019). The origins and consequences of affective polarization in the United States. *Annual Review of Political Science*, *22*, 129–146. https://doi.org/10.1146/annurev-polisci-051117-073034

Iyengar, S., Sood, G., & Lelkes, Y. (2012). Affect, not ideology: A social identity perspective on polarization. *Public Opinion Quarterly*, *76*, 405–431. https://doi.org/10.1093/poq/nfs038

Iyengar, S., & Westwood, S. J. (2014). Fear and loathing across party lines: New evidence on group polarization. *American Journal of Political Science*, *59*(3), 690–707. http://www.jstor.org/stable/24583091

Jaspers, E., Van Zantvliet, P. I., & Willems, J.-W. (2025). How descriptive norms and peer attitudes shape interethnic dating among adolescents in Dutch schools. *Social Inclusion*, *13*, Article 10034. https://doi.org/10.17645/si.10034

Jost, J. T., Baldassarri, D. S., & Druckman, J. N. (2022). Cognitive–motivational mechanisms of political polarization in social-communicative contexts. *Nature Reviews Psychology*, *1*, 560–576. https://doi.org/10.1038/s44159-022-00093-5

Kaba, M., Koyuncu, M., Schneider, S. O., & Sutter, M. (2024). Social norms, political polarization, and vaccination attitudes: Evidence from a survey experiment in Turkey. *European Economic Review*, *168*, Article 104818. https://doi.org/10.1016/j.euroecorev.2024.10481

Kesberg, R., Feddes, A. R., Vogel, E., & Rutjens, B. T. (2025). Associations between perceived societal polarisation and (extreme) non-normative attitudes and behaviour. *Social Inclusion*, *13*, Article 10248. https://doi.org/10.17645/si.10248

Kingzette, J., Druckman, J. N., Klar, S., Krupnikov, Y., Levendusky, M., & Ryan, J. B. (2021). How affective polarization undermines support for democratic norms. *Public Opinion Quarterly*, *85*(2), 663–677. https://doi.org/10.1093/poq/nfab029

Kłodkowski, P., Kossowska, M., & Siewierska, A. (2025). (Micro)Identities in flux: The interplay of polarization and fragmentation in Polish and European politics. *Social Inclusion*, *13*, Article 9996. https://doi.org/10.17645/si.9996

Legros, S., & Cislaghi, B. (2020). Mapping the social-norms literature: An overview of reviews. *Perspectives on Psychological Science*, *15*(1), 62–80. https://doi.org/10.1177/1745691619866455

Levendusky, M. S. (2018). Americans, not partisans: Can priming American national identity reduce affective polarization? *Journal of Politics*, *80*, 59–70. https://doi.org/10.1086/693987

Lutterbach, S., & Beelmann, A. (2025). Youth norm deviation and intolerance: Pathways to polarized political attitudes and behavioral intentions. *Social Inclusion*, *13*, Article 10019. https://doi.org/10.17645/si.10019

McCoy, J., Rahman, T., & Somer, M. S. (2018). Polarization and the global crisis of democracy: Common patterns, dynamics, and pernicious consequences for democratic polities. *American Behavioral Scientist*, *62*, 16–42. https://doi.org/10.1177/0002764218759576

McCoy, J., & Somer, M. (2019). Toward a theory of pernicious polarization and how it harms democracies: Comparative evidence and possible remedies. *Annals of the American Academy of Political and Social Science*, *681*(1), 234–271. https://doi.org/10.1177/0002716218818782

Meleady, R. (2021). "Nudging" intergroup contact: Normative social influences on intergroup contact engagement. *Group Processes and Intergroup Relations*, *24*(7), 1180–1199. https://doi.org/10.1177/13684302211016047

Mudde, C., & Rovira Kaltwasser, C. (2017). *Populism: A very short introduction.* Oxford University Press.

Nijs, T. (2025). Breaking false polarization: How information on descriptive norms mitigates worry rooted in polarization (mis)perceptions. *Social Inclusion*, *13*, Article 9904. https://doi.org/10.17645/si.9904

Orhan, Y. E. (2021). The relationship between affective polarization and democratic backsliding: Comparative evidence. *Democratization*, *29*(4), 714–735. https://doi.org/10.1080/13510347.2021.2008912

Phillips, J. (2022). Affective polarization: Over time, through the generations, and during the lifespan. *Political Behavior*, *44*, 1483–1508. https://doi.org/10.1007/s11109-022-09784-4

Prentice, D., & Paluck, E. L. (2020). Engineering social change using social norms: Lessons from the study of collective action. *Current Opinion in Psychology*, *35*, 138–142. https://doi.org/10.1016/j.copsyc.2020.06.012

Rand, D. G., & Yoeli, E. (2024). Descriptive norms can "backfire" in hyper-polarized contexts. *PNAS Nexus*, *3*(10), Article pgae303. https://doi.org/10.1093/pnasnexus/pgae303

Reiljan, A. (2019). 'Fear and loathing across party lines' (also) in Europe: Affective polarization in European party systems. *European Journal of Political Research*, *59*(2), 376–396. https://doi.org/10.1111/1475-6765.12351

Richters, S., Shani, M., Geyer, L., & Van Zalk, M. H. W. (2025). Differential effects of friendship and school norms on adolescents' defending in cyberbullying situations: A randomized school-based experiment. *Journal of Youth and Adolescence*, *54*(7), 1677–1692. https://doi.org/10.1007/s10964-025-02202-y

Roberts, K. M. (2021). Populism and polarization in comparative perspective: Constitutive, spatial and institutional dimensions. *Government and Opposition*, *57*(4), 680–702. https://doi.org/10.1017/gov.2021.14

Ron, Y., Solomon, J., Halperin, E., & Saguy, T. (2017). Willingness to engage in intergroup contact: A multilevel approach. *Peace and Conflict*, *23*(3), 210–218. https://doi.org/10.1037/pac0000204

Schäfer, L. F., Tausch, N., Bukowski, M., Jaspers, E., Lubbers, M. J., van Zalk, M., Ciordia, C., Potoczek, A., Estevan-Reina, L., Shani, M., Simons, J.-W., Friehs, M.-T., Gurbisz, D., Middendorf, W., Schäfer, S. J., Ullenboom, J., Graf, S., Hjerm, M., Lavest, C., . . . Christ, O. (2024). *Inclusivity norms counter the negative societal effects of opinion polarization across 12 European countries*. OSF. https://doi.org/10.31234/osf.io/9gh6k

Shani, M., Berns, M., Bergen, L., Richters, S., Krämer, K., de Lede, S., & Van Zalk, M. (2025). Longitudinal associations between perceived inclusivity norms and opinion polarization in adolescence. *Social Inclusion*, *13*, Article 10122. https://doi.org/10.17645/si.10122

Shani, M., de Lede, S., Richters, S., Kleuker, M., Middendorf, W., Liedtke, J., Wittola, S., & Van Zalk, M. (2023). A social network intervention to improve adolescents' intergroup tolerance via norms of equality-based respect: The "Together for Tolerance" feasibility study. *International Journal of Developmental Science*, *17*(1/3), 93–110. https://doi.org/10.3233/DEV-230342

Shepherd, H. (2017). The structure of perception: How networks shape ideas of norms. *Sociological Forum*, *32*(1), 72–93. https://doi.org/10.1111/socf.12317

Simons, J. W., Jaspers, E., & Van Tubergen, F. (2025). Socialisation and ethnic majorities' attitudes towards ethnic minorities: A systematic review and meta-analysis of correlational evidence. *Journal of Ethnic and Migration Studies*, *51*(11), 2669–2694. https://doi.org/10.1080/1369183X.2025.2472818

Stavrakakis, Y. (2018). Paradoxes of polarization: Democracy's inherent division and the (anti-) populist challenge. *American Behavioral Scientist*, *62*(1), 43–58. https://doi.org/10.1177/0002764218756924

Suarez Estrada, M., Juarez, Y., & Piña-García, C. A. (2022). Toxic social media: Affective polarization after feminist protests. *Social Media + Society*, *8*(2). https://doi.org/10.1177/20563051221098343

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. 33–47). Brooks/Cole.

Tankard, M. E., & Paluck, E. L. (2016). Norm perception as a vehicle for social change. *Social Issues and Policy Review*, *10*(1), 181–211. https://doi.org/10.1111/sipr.12022

Tausch, N., Birtel, M. D., Górska, P., Bode, S., & Rocha, C. (2024). A post-Brexit intergroup contact intervention reduces affective polarization between Leavers and Remainers short-term. *Communications Psychology*, *2*, Article 95. https://doi.org/10.1038/s44271-024-00146-w

Urbinati, N. (2019). *Me the people: How populism transforms democracy*. Harvard University Press.

Voelkel, J. G., Chu, J., Stagnaro, M., Mernyk, J., Redekopp, C., Pink, S., Druckman, J., Rand, D., & Willer, R. (2022). Interventions reducing affective polarization do not necessarily improve anti-democratic attitudes. *Nature Human Behavior*, *7*, 55–64. https://doi.org/10.1038/s41562-022-01466-9

Wojcieszak, M., Winter, S., & Yu, X. (2020). Social norms and selectivity: Effects of norms of open-mindedness on content selection and affective polarization. *Mass Communication and Society*, *23*(4), 455–483. https://doi.org/10.1080/15205436.2020.1714663

You, Z. T., & Lee, S. W. S. (2024). Explanations of and interventions against affective polarization cannot afford to ignore the power of ingroup norm perception. *PNAS Nexus*, *3*(10), Article pgae286. https://doi.org/10.1093/pnasnexus/pgae286

About the Authors

Miranda J. Lubbers is a full professor of social and cultural anthropology at the Autonomous University of Barcelona, who investigates social networks, social cohesion, and social inclusion.

Marcin Bukowski is a professor at the Jagiellonian University (Institute of Psychology) and a member of the Centre for Social Cognitive Studies in Kraków, Poland. His research interests focus mainly on motivated social cognition and the psychology of intergroup relations.

Oliver Christ is a full professor at the FernUniversität in Hagen (Faculty of Psychology), Germany. His research interests focus mainly on the social psychology of intergroup relations.

Eva Jaspers is a full professor of empirical and theoretical sociology at Utrecht University. Her research is on gender, sexual, and racial/ethnic diversity from an interactional or network perspective. She aims to understand how people's interactions and relations with (dis)similar others shape their patterns of behavior and their outcomes at school, work, and at home.

Maarten van Zalk is a full professor of developmental psychology at Osnabrück University who focuses on adolescent development of intergroup relations, social relationships, and antisocial behavior.

cogitatio

Testing the Robustness of the Association Between Personal Respect Norms and Tolerance in Polarized Contexts

Lucía Estevan-Reina [1,†], **Laura Frederica Schäfer** [2,†], **Wilma Middendorf** [3], **Marcin Bukowski** [4], **Maarten van Zalk** [3], **and Oliver Christ** [2]

[1] Department of Social Psychology, University of Granada, Spain
[2] Department of Psychological Methods and Evaluation, FernUniversität in Hagen, Germany
[3] Department of Developmental Psychology, University of Osnabrück, Germany
[4] Institute of Psychology, Jagiellonian University, Poland

[†] These authors have contributed equally to this work and share first authorship.

Correspondence: Oliver Christ (oliver.christ@fernuni-hagen.de)

Submitted: 31 January 2025 **Accepted:** 9 May 2025 **Published:** 23 July 2025

Issue: This article is part of the issue "The Impact of Social Norms on Cohesion and (De)Polarization" edited by Miranda Lubbers (Autonomous University of Barcelona), Marcin Bukowski (Jagiellonian University), Oliver Christ (FernUniversität in Hagen), Eva Jaspers (University of Utrecht), and Maarten van Zalk (University of Osnabrück), fully open access at https://doi.org/10.17645/si.i438

Abstract

Societies worldwide are challenged by heated debates around important societal topics like migration policies, gender equality, transgender rights, and climate change. These debates are perceived as highly polarized thereby increasing intolerance toward opposing opinions. Previous research has shown that respecting "disapproved others" as equals might foster tolerance, even in polarized contexts. Yet, an empirical test to establish whether the relationship link between respect and tolerance toward opposing others is still observable in the case of extreme opinions, strong disapproval of opposing opinions, and even strong perceived threats from opposing others, is still missing. In our research, we will test whether the strength of the association between personal respect norms and the tolerance of opposing opinions depends on the extremity of one's own opinion, the strength of disapproval of the opposing opinion, and the perceived threat from the out-group. Results based on survey data from more than 12,000 respondents from 12 European countries reveal that the association between personal respect norms and tolerance is unaffected by extremity, strength of disapproval, and perceived threat. The pattern of results is replicated with few exceptions across all 12 countries and six different controversial social topics. This is held in most cases even when considering differences in political views. We discuss the implications of our findings, their robustness, and the potential limits of the respect–tolerance link.

Keywords

norms; polarization; respect; tolerance

---

1. Introduction

In many societies, discourses on important social topics are perceived as becoming increasingly polarized (Koudenburg & Kashima, 2022). Research on perceived polarization highlights its negative consequences even beyond those of actual polarization (Lelkes, 2016; Westfall et al., 2015; Wilson et al., 2020), especially when it is accompanied by a strong aversion toward those who hold opposing opinions (i.e., affective polarization; see, e.g., Iyengar et al., 2012, 2019) and an "us versus them" mentality (Jost et al., 2022). Some of the negative consequences include increased intolerance, avoidance of opposing groups, and open hostility (e.g., Berntzen et al., 2024). In the face of perceived polarization and the growing dislike and distrust toward individuals with opposing opinions, a promising strategy is to foster tolerance by cultivating respect for "disapproved others" (i.e., grounded in the recognition of others as equals; Renger & Simon, 2011; Simon, 2023; Verkuyten et al., 2023). Studies indicate that respect for groups whose cultural practices or political views are disapproved can enhance tolerance toward these practices or views (e.g., Eschert & Simon, 2019; Simon, 2023; Velthuis et al., 2021). Consequently, fostering respect appears to be an effective way to increase tolerance in polarized contexts, particularly when respect is perceived to be a generalized principle in society (i.e., a social norm) and applied to all societal groups and individuals, including those with differing opinions. Initial empirical evidence supports this, showing that respect norms are associated with tolerance toward disapproved and disliked others (Schäfer et al., 2024). However, the role of personal norms (internalized norms) of respect in promoting tolerance has not yet been evaluated and research on potential boundary conditions for the respect norms–tolerance link is limited. Thus, it remains unclear how respect norms are manifest in those with extreme positions (i.e., when disapproval of another opinion is stronger, or when one's own opinion is more extreme) or who feel threatened by those with opposing opinions. For instance, studies suggest that intolerance and negative behaviours toward groups and individuals with differing opinions are more likely when opinions are more extreme (Bos et al., 2023; Van Boven et al., 2012; Westfall et al., 2015), or when individuals feel threatened by those with opposing opinions (Carriere et al., 2022). To fully understand the potential of respect norms in fostering tolerance in polarized contexts, it is crucial to investigate potential boundary conditions. This study addresses the following general research question: Does the relationship between personal respect norms and tolerance persists when individuals hold extreme positions on a controversial social topic and perceive a greater threat from those holding opposing opinions?

1.1. The Respect–Tolerance Link as a Means to Counter Polarization

Tolerance is often regarded as a fundamental cornerstone of democratic societies, enabling individuals and groups with differing views, opinions, and practices to coexist peacefully (Verkuyten, 2023). Importantly, tolerance does not require abandoning one's own beliefs or ceasing to disapprove of opposing opinions. Instead, it involves accepting differing beliefs, opinions, and practices while maintaining one's disapproval. This nuanced understanding of tolerance arises from balancing reasons for disapproval with reasons to accept dissenting opinions, beliefs, or practices (e.g., Verkuyten et al., 2023). When the reasons for acceptance outweigh those for rejection, tolerance becomes possible (Gibson, 1989; Sullivan et al., 1979; Verkuyten, 2023). Respecting others as equal fellow citizens, which is a fundamental component of democratic societies, serves as one promising counterweight (i.e., the respect conception of tolerance; Forst, 2013; Scanlon, 2003). Indeed, recent research shows that respect is a powerful predictor of tolerance: Individuals who respect others as equals, despite their disapproval of them, are more likely to grant them the

rights to live according to their chosen way of life (Simon, 2023; Simon & Schaefer, 2016; Zitzmann et al., 2022). Moreover, research suggests that the respect conception of tolerance is particularly promising in highly diverse societies, proving more effective than other forms of tolerance, such as coexistence tolerance (Velthuis et al., 2021). For a discussion of different conceptions of tolerance see Forst (2013), Hjerm et al. (2020), and Verkuyten and Yogeeswaran (2017). There is also initial evidence that respecting others as equals helps to overcome bias in polarized political contexts. Eschert and Simon (2019) experimentally showed that inducing respect can reduce the negative evaluation of out-group arguments on contested topics. Thus, promoting respect seems to be an effective strategy for fostering tolerance in polarized contexts, especially when respect is perceived as a generalized principle in society (i.e., a social norm) applied to all societal groups and individuals with varying perspectives (Schäfer et al., 2024).

1.2. Respect Norms and Tolerance

A substantial body of research has demonstrated that perceived social norms play a key role in shaping individual behaviour, serving as the informal rules that guide our daily lives (Bicchieri et al., 2023; Tankard & Paluck, 2016). Perceiving what others approve of (injunctive norms) or commonly do (descriptive norms) can motivate individuals to act similarly (Cialdini et al., 1990; Miller & Prentice, 2016; Nolan, 2021). Research has shown that perceived social norms help to improve intergroup relations such as reducing hate speech (Bilewicz et al., 2021), promoting positive attitudes towards out-groups, greater appreciation of diversity (Murrar et al., 2020), willingness for intergroup contact (Meleady, 2021), and interest in cross-group friendships (Tropp et al., 2014). There is also evidence that social norms of tolerance can foster political tolerance (i.e., the willingness to permit the expression of ideas or interests that one opposes; Neuner & Ramirez, 2023). Moreover, in a recent cross-European survey study, Schäfer et al. (2024) showed that inclusivity norms—norms that define how to approach members of society with differing opinions by emphasizing respect, dialogue, and unity—are associated with greater tolerance, increased willingness to collaborate, and reduced tendencies to avoid individuals with opposing opinions. These findings provide the first evidence for the potential of respect norms to avoid the negative consequences of increased levels of perceived polarization in societies. Further experimental evidence shows that respect norms increase tolerance toward LGBTQ+ rights in Poland (Estevan-Reina et al., 2024).

In our research, we specifically focus on personal norms (Bicchieri, 2017), which are defined as social norms that are internalized—that is, personal beliefs that align with societal standards. Personal norms thereby become an independent motivation to act (Thøgersen, 2006) and form a significant part of an individual's identity (Bar-On & Lamm, 2023). There is evidence that such (internalized) personal norms are strong predictors of behaviour (e.g., Cialdini & Jacobson, 2021). In the present research, we measure interindividual differences in personal respect norms. In that regard, personal respect norms represent individuals' support for the belief that all members of society should be treated as equals. We assume that personal respect norms are associated with tolerance towards individuals and groups with differing and disapproved opinions on important social topics.

1.3. Potential Boundary Conditions for the Link Between Personal Respect Norms and Tolerance

Individuals vary not only in their opinions on controversial social topics but also in the extremity of those opinions and the strength of disapproval of opposing opinions on such topics. Likewise, individuals can

perceive those with opposing opinions as more or less threatening. Research has shown that the extremity of an opinion on social topics (i.e., extreme opinion and/or strong disapproval of differing opinions) and perceived threat directly influence the approach toward opposing others (e.g., with more negative attitudes, more avoidance) and whether they are tolerated or not. For instance, Bos et al. (2023) showed that respondents with more extreme attitudes are more intolerant towards those with opposing opinions. Westfall et al. (2015) found that people perceive greater political polarization when they hold extreme partisan attitudes aligned with their own identity due to a polarization projection (Van Boven et al., 2012). Likewise, there is ample evidence showing that perceived intergroup threat is associated with more negative attitudes (Riek et al., 2006), the restriction of human rights (see Carriere et al., 2022), and less tolerance (e.g., Chanley, 1994; Crawford & Pilanski, 2014; Haas & Cunningham, 2014; Skitka et al., 2004).

Therefore, extreme positions concerning a controversial social topic and perceived threats from those with opposing opinions might also influence the extent to which personal respect norms translate into tolerance toward those with opposing opinions. In cases of extreme positions and higher perceived threats, reasons to reject opposing opinions might outweigh reasons to accept them, making tolerance less likely (Verkuyten et al., 2023). Thus, extremity of the "own opinion," strength of disapproval, and perceived threat might moderate the link between personal respect norms and tolerance. It is therefore important to test whether the association between personal respect norms and tolerance towards those with an opposing opinion on a controversial social topic is weakened or even absent when individuals strongly disapprove of those with opposing opinions, have an extreme position on the respective social topic, and feel threatened by those with differing opinion.

1.4. The Role of Political Orientation and Opinions Across Topics

We will also explore potential differences in the strength of association between personal respect norms and tolerance depending on the political orientation of individuals as well as their opinion on a controversial social topic (i.e., progressive or conservative). Several scholars argue that liberals and conservatives are equally intolerant of those with ideologically dissimilar views or those of whom they disapprove (Brandt et al., 2014; Chambers et al., 2013; Crawford, 2014). However, there is also evidence that liberal (or left-leaning) individuals tend to emphasize more strongly the importance of equality, diversity, and tolerance of differences compared to conservative (or right-leaning) individuals (Jost, 2017) and, therefore, tend to be more tolerant of differences (Badaan & Jost, 2020).

Moreover, the level of perceived polarization varies between controversial social topics. For instance, Herold et al., (2023) found substantial variation in the level of affective polarization based on survey data from Europe, with high levels of polarization for the topics of "climate change" and "immigration" and comparably low levels of polarization for the topics of "gender equality in society" and "social benefits and their financing."

For a more exploratory purpose, we will also examine the potential influence of political orientation and type of opinion (i.e., progressive or conservative) on a given controversial topic on the personal respect norms–tolerance link alongside the hypothesized moderating influence of extremity of the own opinion, strength of disapproval of the opposite opinion, and perceived threat.

1.5. Aim of the Study/Research Questions

To fully understand the potential of respect norms in fostering tolerance in contexts that are perceived to be polarized, it is crucial to investigate potential boundary conditions of the respect norms–tolerance link (Busse et al., 2017). For this reason, in this preregistered study (see Data Availability, for more information), we will test our research questions using survey data from more than 12,000 respondents from 12 European countries. Specifically, we will test whether the strength of the association between personal respect norms and tolerance of opposing opinions on different controversial social topics is dependent on the extremity of one's own opinion (RQ1), the strength of disapproval of the opposing opinion (RQ2), and the perceived threat from those with opposing opinions (RQ3). Additionally, we will explore whether the pattern of results is comparable for left-leaning and right-leaning individuals (RQ4.1) as well as for those holding progressive (e.g., in favour of accepting more refugees in the country) and conservative opinions (e.g., against accepting more refugees in the country) as a moderator (RQ4.2) across topics.

2. Methods

2.1. Respondents

This study uses data from a large cross-national survey (Schäfer et al., 2024) conducted between November 9 and December 6, 2023, as part of a research project concerning social norms in polarized contexts (INCLUSIVITY), with approximately 1,000 respondents from each of the 12 European countries, including countries from Northern (Finland, Sweden), Western (France, Germany, the Netherlands, and UK), Southern (Italy, Greece, and Spain), and Eastern Europe (Czechia, Hungary, and Poland). Data were collected with web-based questionnaires by an established German survey company with access to national data online panels across Europe. The samples were quoted to approximate the demographic composition based on census data of each country in terms of age, education, and gender (for more details of the sample compositions see the Supplementary File, Table S1).

2.2. Procedure

Respondents were first presented with instructions emphasizing that the following questions pertain to all individuals living in their country, encompassing people from diverse social groups based on cultural or ethnic backgrounds, as well as those who shared opinions on controversial social topics. Thereafter, respondents' personal respect norms were measured. Tolerance, extremity of one's own opinion, and strength of disapproval of the opposite opinion were assessed using an adapted version of the least-liked-group approach focusing on groups with opposing opinions (Gibson, 1992; Skitka et al., 2013). This approach ensured that respondents had an opinion on the topic they were presented with, identified to some degree with a group linked to the topic, and exhibited disapproval towards individuals holding the opposite opinion. The procedure consisted of three steps. First, from a list of six controversial social topics (i.e., climate change, migration and refugee policies, gender equality, transgender rights, Covid-19 vaccination, and meat consumption) respondents chose the one on which they felt they had the firmest stance. The topics provided were selected based on previous research on groups with opposing opinions and polarization in Europe (Agostini & van Zomeren, 2021; Herold et al., 2023). After selecting a topic (e.g., asylum and migration policy), respondents stated their opinion related to that topic as being either in

favour (e.g., accepting more refugees in the country) or against (e.g., not accepting more refugees in the country). Subsequently, respondents indicated the extremity of their opinion, the strength of their disapproval of the opposite opinion, and their perceived threat from people with opposing opinions on the contested topic. Thereafter, respondents' tolerance was assessed. Finally, political orientation was measured along with sociodemographic information at the end of the survey.

2.3. Measures

The main independent variable, *personal respect norms*, was assessed using three items adapted from previous research: "Everyone should always be treated as a human being of equal worth"; "All in all, everyone should be treated equally" (Renger et al., 2017); "Everyone should be recognized as a fellow citizen of equal worth" (Simon & Schaefer, 2018). The items were measured on 7-point Likert scales ranging from 1 (*strongly disagree*) to 7 (*strongly agree*). Omega reliability coefficients for these items ranged from .77 to .90.

The moderators were measured with single items in the case of extremity of opinion and strength of disapproval and with multiple items in the case of perceived threat. *Extremity of the own opinion* was measured after respondents selected a topic and indicated their opinion. Using one item adapted from Skitka et al. (2013), respondents answered the question, "How strongly do you hold this opinion?" on a 7-point Likert scale ranging from 1 (*very weakly*) to 7 (*very strongly*). *Strength of disapproval* was assessed with one item on a 7-point Likert scale ranging from 1 (*very weak*) to 7 (*very strong*). Respondents were asked: "How strong is your disapproval regarding the opinion that [opposite opinion]?" *Perceived threat* from individuals holding an opposing opinion was measured across four dimensions: *sociotropic threat* (i.e., perceived threat to the normative social order); *egocentric threat* (i.e., perceived threat to an individual's personal freedom); *threat to safety*; and *threat to autonomy*. The first three dimensions were based on Gibson et al. (2020), while the items measuring perceived *threat to safety* were adapted from Doosje et al. (2012). Sociotropic threat, egocentric threat, and perceived threat to safety were measured on a 7-point Likert scale ranging from 1 (*does not apply at all*) to 7 (*applies completely*). The two items assessing *sociotropic threat* read: "People who [opposite opinion] are dangerous for [country] society in general" and "People who [opposite opinion] are dangerous for the normal lives of people in [country]." Two items assessing *egocentric threat*: "People who [opposite opinion] reduce my personal political freedom" and "People who [opposite opinion] would affect my personal safety if they gained power," and two items measuring *threat to safety*: "People who [opposite opinion] are prepared to use violence against other people to achieve something they consider very important" and "People who [opposite opinion] are prepared to disrupt public order to achieve something they consider very important." *Threat to autonomy* was assessed on a scale ranging from 1 (*not at all likely*) to 7 (*very likely*) using the following two items: "How likely is it that people who [opposite opinion] will gain a lot of power in [country] society?" and "How likely is it that people who [opposite opinion] will affect your or your family's quality of life?" To measure perceived threat from the out-group, we combined the eight items derived from the four subscales described above. Omega coefficients for the composite scale ranged from .72 to .92. Additionally, we measured *political orientation* using an 11-point left–right scale ranging from 0 (left) to 10 (right) and that read as follows: "People often talk about 'left' and 'right' in politics. How would you classify your basic political stance?"

The main dependent variable, *tolerance*, was assessed using four items adapted from Simon et al. (2019) and had already been used in a recent work/publication by Schäfer et al. (2024). Respondents indicated their

disagreement or agreement to the following four items on 7-point Likert scales ranging from 1 (*strongly disagree*) to 7 (*strongly agree*): "These people should be able to stick to their opinion, even if others try to persuade them otherwise," "These people should be allowed to speak their mind freely and openly"; "These people should be allowed to meet in public spaces and give speeches"; "These people should be given the chance to pursue their interests just like others." Omega reliability coefficients for these items ranged from .78 to .87, indicating good internal consistency.

3. Results

3.1. Descriptives

The ranges of means, standard deviations, and bivariate correlations for all measures and countries are presented in Table 1 (for descriptives and correlations for each country see the Supplementary File, Tables S2–S13). Across all countries, and consistent with previous research (Simon & Schaefer, 2016, 2018), we find positive and significant correlations between personal respect norms and tolerance, with correlation coefficients (r) ranging from .15 to .37.

Bivariate correlations between personal respect norms and the moderators are generally low, indicating that these variables are mostly unrelated. Specifically, correlations between extremity of the own opinion and personal respect norms range from $r = -0.02$ to $r = 0.16$. Strength of disapproval is largely unrelated to respect norms ($r = 0.00$ to $r = -0.10$), whereas correlations between perceived threat and personal respect norms are generally significant, ranging from $r = -0.06$ to $r = -0.25$. The low correlations show the advantage of using the adapted version of the least-liked-group approach: Individuals with high and low personal respect norms reported comparable scores on how strongly they held their views, how much they disapproved of opposing views, and how threatened they felt, thus, making comparisons between these groups more meaningful.

Bivariate correlations between the moderators and tolerance are mostly negative. Extremity of the own opinion is negatively and significantly correlated with tolerance in six out of twelve countries (Czechia, Finland, Greece, Italy, Spain, and the UK; $r = -.08$ to $r = -.23$), but shows no significant correlations in the remaining six countries (France, Germany, Hungary, the Netherlands, Poland, and Sweden). Strength of disapproval shows consistently negative and significant correlations with tolerance across all countries ($r = -.10$ to $r = -.44$), with the exception of the Netherlands. Similarly, perceived threat is significantly and negatively related to tolerance in all countries ($r = -.20$ to $r = -.44$).

With regard to political orientation, bivariate correlations with personal respect norms are generally significant and negative. This indicates that the individuals who lean more to the left end of the political orientation scale tend to endorse personal respect norms more strongly ($r = -.04$ to $r = -.27$). Correlations between political orientation and tolerance are mostly non-significant. However, we find significantly negative correlations in France, Germany, and Hungary ($r = -.07$ to $r = -.13$), but a significantly positive correlation in Spain ($r = .08$).

Table 1. Range of means and standard deviations for all countries.

Variable	*M*	*SD*	1	2	3	4	5
1. Personal respect norms	5.76–6.32	0.89–1.33					
2. Disapproval	4.69–5.32	1.58–1.96	.00 – −.10**				
3. Extremity	5.81–6.23	0.97–1.33	−.02–.16**	.18**–.48**			
4. Threat	3.94–4.47	1.31–1.57	−.03 – −.25**	.10*–.45**	.17**–.32**		
5. Political orientation[1]	5.56–6.43	1.85–2.74	−.04 – −.27**	.01 – −.15**	.00 – −.14**	−.01–.17**	
6. Tolerance	4.69–5.11	1.18–1.40	.15**–. 37**	.04 – −.44**	−.01 – −.23**	−.20** – −.44**	.00 – −.13**

Notes: *M* and *SD* are used to represent mean and standard deviation, respectively; intervals represent the minimum and maximum values; [1] 10-point scale (1 = left to 11 = right); * $p < .05$. ** $p < .01$.

3.2. Main Analysis

To test RQ1–RQ3, we conducted multiple regression analyses for each country sample (Czechia, Finland, France, Germany, Greece, Hungary, Italy, the Netherlands, Poland, Spain, Sweden, and the UK) using the *dplyr* package in R (RStudio version 2024.09.1+394). In all regression analyses reported, variables used to create interaction terms were mean-centred. We tested the effect of each moderator (extremity of the own opinion, strength of disapproval of the other opinion, and perceived out-group threat) in separate analyses. Specifically, each model included personal respect norms, each of the aforementioned moderators, and their interaction terms as predictors, allowing us to examine whether the association between personal respect norms and the tolerance varies across levels of each moderator.

For *extremity of the own opinion* (RQ1), there were no significant two-way interactions with personal respect norms in 11 out of the 12 countries (see the Supplementary File, Tables S14–S25), suggesting that the association between personal respect norms and tolerance is comparable for individuals holding more extreme opinions compared to those holding less extreme opinions. Only in Poland did a significant two-way interaction with personal respect norms on tolerance emerge ($B = .06$, $SE = .03$, $p = .03$), with the association being stronger for individuals holding more extreme opinions compared to those holding less extreme opinions (for simple slopes see the Supplementary File, Figure S1).

When we included *strength of disapproval* of the opposite opinion as a moderator (RQ2), there were no significant two-way interactions with personal respect norms in any of the countries (see the Supplementary File, Tables S14–S25). Again, this suggests that the link between personal respect norms and tolerance is not dependent on individuals' strength of disapproval.

The inclusion of *perceived threat* (RQ3) revealed no significant two-way interactions across countries (see the Supplementary File, Tables S14–S25), indicating that perceived threat does not affect the association between personal respect norms and tolerance in most cases. There is an exception in the case of the UK ($B = −.09$, $SE = .02$, $p < .001$; for more details see the Supplementary File, Table S25), with the relationship being stronger for individuals perceiving less threat compared to those who perceive more (for simple slopes see the Supplementary File, Figure S2).

Overall, results indicate that the association between personal respect norms and tolerance is robust and comparable for respondents differing in extremity of the own opinion, strength of disapproval, and perceived threat. To test the robustness of results of the main analysis, we then included the three-way interaction between personal respect norms, political orientation, and each of the moderators separately.

3.3. Exploratory Analyses

To explore its role more closely, we included political orientation as a moderator to test whether the strength of the relationship between personal respect norms and tolerance depends on political orientation direction (left–right leaning; RQ4.1). The results indicate mixed evidence (see the Supplementary File, Tables S26–S37). While there was no significant two-way interaction between political orientation and personal respect norms in 7 out of the 12 countries (see the Supplementary File, Tables S14–S25), significant two-way interactions between political orientation and personal respect norms emerged in five out of the 12 (i.e., Finland, Germany, Greece, Italy, and Poland), $Bs \geq .04$, $SEs \leq .02$, $ps \leq .01$ (for simple slopes see the Supplementary File, Figure S3–S7). These findings suggest that in some countries, the association between personal respect norms and tolerance is stronger for politically right-leaning individuals. To test the robustness of the results of the main analysis, we then included the three-way interaction between personal respect norms, political orientation, and each of the moderators separately (strength of disapproval, extremity of the own opinion, and perceived threat). Including political orientation as a moderator introduces certain limitations, making the interpretation complicated. First, our analysis focused solely on the *direction* of political orientation, without accounting for its *extremity*. Second, using political orientation as a moderator in conjunction with the adapted version of the least-liked approach may constrain the interpretation of our results. In the interest of simplicity and clarity, we chose not to include the full results in the main manuscript. However, as this was a preregistered research question, the complete analyses are available in the Supplementary File (Tables S26–S37 and Figures S10–S18).

Additionally, we examined whether the strength of relationships varied across topics by grouping respondents based on the topic they chose (RQ4.2), rather than by country. Specifically, we explored whether the strength of the relationship between personal respect norms and tolerance, as well as the strength of each moderator, depended on respondents' opinions on the relevant societal issue within each topic (i.e., climate change, migration policies, gender equality, transgender rights, Covid-19 vaccination, and meat consumption; for more details on topic selection and distribution of opinions on the selected topics for each country see the Supplementary File, Table S50).

To do so, we repeated the analyses conducted for political orientation but included respondents' opinions on the respective topic as an additional predictor instead. We also reran all analyses reported in the main analysis section, as the analytic sample changed. In the main analyses, we conducted separate regression analyses for each country, whereas in these exploratory analyses, we sampled respondents by topic, combining individuals from different countries. Overall, the pattern of results remained unchanged.

The analyses revealed no significant interactions between personal respect norms and either *strength of disapproval* or *extremity of the own opinion* for any of the topic samples (see the Supplementary File, Tables S38–S43). For *perceived threat*, there were no significant two-way interactions in the case of migration policies, transgender rights, Covid-19 vaccination, and meat consumption. However, significant

two-way interactions emerged in the case of climate change, $B = -.04$, $SE = .01$, $p = .01$, and gender equality, $B = -.04$, $SE = .02$, $p = .02$ (for simple slopes see the Supplementary File, Figure S19 and S20).

We then included the three-way interaction between personal respect norms, opinion on the topic (i.e, progressive versus conservative), and each moderator separately. Most of the three-way interactions were non-significant, including those between opinion on the topic, personal respect norms, and *extremity of the own opinion* or *perceived threat*, respectively, across all topic groups. The interaction between opinion on the topic, personal respect norms, and *strength of disapproval* became significant for three out of the six topic groups (i.e., climate change, migration and asylum policy, and transgender rights), $Bs \geq .07$, $SEs \leq .06$, $ps \leq .05$. These findings suggest that the interaction between strength of disapproval and personal respect varies depending on individuals' opinions on these topics (for simple slopes see the Supplementary File, Figures S21–S23; see regression Tables S44–S49).

Specifically, for *climate change*, among individuals who believe it is man-made, the association between personal respect norms and tolerance tends to be weaker for those with more, than for those with less disapproval. For individuals who believe climate change is not man-made, the association is weaker for those holding less disapproval.

A similar pattern emerged for *migration and asylum policy*. Among those who are in favour of accepting more refugees in their country, the association between personal respect and tolerance was weaker for those holding more disapproval than those holding less. The pattern reverses for individuals who are against accepting more refugees in their country: The association between personal respect norms and tolerance was weaker for those holding less disapproval than for those holding more.

In the case of *transgender rights*, individuals who are in favour of gender self-determination showed a stronger association between personal respect norms and tolerance when holding weaker disapproval. The pattern was reversed for those opposing gender self-determination.

4. Discussion

In this article, we tested whether personal respect norms translate into tolerance toward those with opposing opinions, even in the case of extreme positions on controversial social topics (i.e., strong disapproval of the opposing opinion or extreme opinion on a social topic) and perceived threat toward those with opposing opinions. While previous research has shown that respecting others as different equals or perceiving social norms that emphasize respect are associated with tolerating disapproved practices and views (e.g., Eschert & Simon, 2019; Schäfer et al., 2024; Simon, 2023; Velthuis et al., 2021), potential boundary conditions have not yet been tested. In cases of extreme position and higher perceived threats that are more likely to be observed in polarized settings, reasons to reject opposing opinions might outweigh reasons to accept these opinions, making tolerance less likely (Verkuyten et al., 2023).

Our results, based on survey data including 12,000 respondents from 12 European countries clearly show that the respect–tolerance link holds even among those who endorse extreme positions on a controversial social topic (i.e., holding an extreme opinion or strong disapproval of different opinions; in line with Schäfer et al., 2024), and feel threatened by those with opposing opinions. This general pattern of results was also

observable among left—and right-leaning respondents (for details see the Supplementary File) as well as for differing opinions across the social topics examined. Thus, our results indicate that the association of personal respect norms with tolerance is relatively robust and might therefore serve as an approach to foster tolerance even in contexts of perceived polarization that are often accompanied by an increase in mutual dislike and distrust. However, this does not mean that we should discourage efforts to reduce extreme positions on controversial social issues or perceived threats toward those with differing opinions, as these factors are often related to lower levels of tolerance (See Table 1). On the contrary, further research in this area remains essential. Nevertheless, considering our findings, we propose that promoting norms of respect may serve as a complementary approach—and in some cases, a more practical and effective one, since changing norms seems to be easier than attitudes (Prentice & Paluck, 2020).

Crucially, our claim is not that everything should be unconditionally tolerated. Beliefs and practices that violate general principles and moral values (Verkuyten et al., 2020) present valid reasons for witholding tolerance, even when individuals recognize the equal rights of others. This might also explain why the association between personal respect norms and tolerance is weaker in some cases, as our exploratory analyses show when taking political orientation and opinions on controversial social topics into account. For instance, among left-leaning individuals in some of the countries examined (see the Supplementary File, Figures S10–S18 and Tables S26–S37) or among individuals with progressive opinions on controversial social topics (e.g., supporting the idea that climate change is man-made, favouring migration and asylum policies, or endorsing self-determination laws), the association between personal respect norms and tolerance is weaker for those who hold more extreme opinions, show higher levels of disapproval toward opposing opinions, and feel more threatened by those with opposing opinions. These results show that personal respect norms do not always translate to the same extent into tolerance of opposing opinions. This might be an outcome of reflective reasoning, in which individuals weigh reasons for rejection (e.g., potential harm for immigrants) against reasons for acceptance (deliberative intolerance; see Verkuyten et al., 2020). However, we cannot rule out that in such cases other factors can lead people to express less tolerance (or intolerance), such as stronger negative feelings towards those with opposing opinions or immediate intuitions and emotions. These forms of less tolerance (or intolerance) have been referred to as "prejudicial" (in cases of negative out-group feelings) and "intuitive" (in cases of negative feelings and intuitions) intolerance by Verkuyten et al. (2020).

We acknowledge that our research has several limitations. First, the results are based on cross-sectional data, which inherently limits the ability to draw causal conclusions. Second, our reliance on self-reported measures raises concerns about whether reported attitudes truly translate into tolerant behaviour in real-life situations (e.g., Dixon et al., 2017). Additionally, while the adaptation of the least-liked-group approach that we used offers notable strengths (ensuring that the topic chosen by the respondents was relevant to them, that they had an established opinion on the issue, and disapprove of the opposing opinion, enabling us to measure tolerance toward disapproved others for all respondents), it also introduces complexities in interpreting some of the findings. For instance, the selection of topics may be influenced by an individual's political orientation, which in turn affects the level of disapproval toward opposing opinions, the extremity of one's own opinion, and the perceived threat from opposing groups. Finally, although our samples are large enough to be considered representative (approximately 1,000 respondents per country), the number does not guarantee full representativeness, and participants are drawn from Western, educated, industrialized, rich, and democratic (WEIRD) societies. This limits the generalizability of our findings to other cultural contexts.

5. Conclusion

Despite the limitations, we believe this work makes a valuable contribution to the literature. Personal respect norms seem to translate into tolerance even in the case of extreme positions on a certain social issue or in the case of higher perceived threats from those with opposing opinions. For the most part, this pattern of results holds for left—as well as right-leaning individuals (although future research should examine more exhaustively the role of political orientation—direction and extremity—as a potential moderator) and among people who have progressive and conservative opinions about different controversial social topics. Lastly, but importantly, our results are based on large samples from 12 European countries with different degrees of (affective) polarization (Herold et al., 2023) and, with few exceptions, the pattern of results replicates across all the countries and six different controversial social topics. To conclude, this work shows that personal respect norms help to tolerate those with disapproved opposing opinions on important social issues, granting them the right to stick to their opinion, speak up openly in public, and pursue their interests given that all citizens have the right to do so in a functioning democratic society.

Conflict of Interests

The authors declare no conflict of interests.

Data Availability

Data have been collected as part of a larger research project and, thus, cannot be shared before the project has been officially ended. Data can be requested from corresponding authors until then and will be made publicly available in 2026. For more information on the preregistered study, including the research questions and methodology, please see the preregistration on the Open Science Framework: https://osf.io/ej8fd The study uses survey data from over 12,000 respondents across 12 European countries. A detailed psychometric report, comprehensive descriptions of all measures, and the full analysis script are available at: https://osf.io/afxcg

Supplementary Material

Supplementary material for this article is available online in the format provided by the author (unedited).

References

Agostini, M., & van Zomeren, M. (2021). Toward a comprehensive and potentially cross-cultural model of why people engage in collective action: A quantitative research synthesis of four motivations and structural constraints. *Psychological Bulletin*, *147*(7), 667–700. https://doi.org/10.1037/bul0000256

Badaan, V., & Jost, J. T. (2020). Conceptual, empirical, and practical problems with the claim that intolerance, prejudice, and discrimination are equivalent on the political left and right. *Current Opinion in Behavioral Sciences*, *34*, 229–238. https://doi.org/10.1016/j.cobeha.2020.07.007

Bar-On, K., & Lamm, E. (2023). The interplay of social identity and norm psychology in the evolution of human groups. *Philosophical Transactions of the Royal Society of London B*, *378*(1872), Article 20210412. https://doi.org/10.1098/rstb.2021.0412

Berntzen, L. E., Kelsall, H., & Harteveld, E. (2024). Consequences of affective polarization: Avoidance, intolerance and support for violence in the United Kingdom and Norway. *European Journal of Political Research*, *63*, 927–949. https://doi.org/10.1111/1475-6765.12623

Bicchieri, C. (2017). *Norms in the wild. How to diagnose, measure, and change social norms*. Oxford University Press.

Bicchieri, C., Dimant, E., Gelfand, M., & Sonderegger, S. (2023). Social norms and behavior change: The interdisciplinary research frontier. *Journal of Economic Behavior & Organization*, *205*, A4–A7. https://doi.org/10.1016/j.jebo.2022.11.007

Bilewicz, M., Tempska, P., Leliwa, G., Dowgiałło, M., Tańska, M., Urbaniak, R., & Wroczyński, M. (2021). Artificial intelligence against hate: Intervention reducing verbal aggression in the social network environment. *Aggressive Behavior*, *47*(3), 260–266. https://doi.org/10.1002/ab.21948

Bos, L., Wichgers, L., & van Spanje, J. (2023). Are populists politically intolerant? Citizens' populist attitudes and tolerance of various political antagonists. *Political Studies*, *71*(3), 851–868. https://doi.org/10.1177/00323217211004592

Brandt, M. J., Reyna, C., Chambers, J. R., Crawford, J. T., & Wetherell, G. (2014). The ideological-conflict hypothesis: Intolerance among both liberals and conservatives. *Current Directions in Psychological Science*, *23*(1), 27–34. https://doi.org/10.1177/0963721413510932

Busse, C., Kach, A. P., & Wagner, S. M. (2017). Boundary conditions: What they are, how to explore them, why we need them, and when to consider them. *Organizational Research Methods*, *20*(4), 574–609. https://doi.org/10.1177/1094428116641191

Carriere, K. R., Hallahan, A., & Moghaddam, F. M. (2022). The effect of perceived threat on human rights: A meta-analysis. *Group Processes & Intergroup Relations*, *25*(1), 247–279. https://doi.org/10.1177/1368430220962563

Chambers, J. R., Schlenker, B. R., & Collisson, B. (2013). Ideology and prejudice: The role of value conflicts. *Psychological Science*, *24*(2), 140–149. https://doi.org/10.1177/0956797612447820

Chanley, V. (1994). Commitment to political tolerance: Situational and activity-based differences. *Political Behavior*, *16*(3), 343–363. https://doi.org/10.1007/BF01498955

Cialdini, R. B., & Jacobson, R. P. (2021). Influences of social norms on climate change-related behaviors. *Current Opinion in Behavioral Sciences*, *42*, 1–8. https://doi.org/10.1016/j.cobeha.2021.01.005

Cialdini, R. B., Reno, R. R., & Kallgren, C. A. (1990). A focus theory of normative conduct: Recycling the concept of norms to reduce littering in public places. *Journal of Personality and Social Psychology*, *58*(6), 1015–1026. https://doi.org/10.1037/0022-3514.58.6.1015

Crawford, J. T. (2014). Ideological symmetries and asymmetries in political intolerance and prejudice toward political activist groups. *Journal of Experimental Social Psychology*, *55*, 284–298. https://doi.org/10.1016/j.jesp.2014.08.002

Crawford, J. T., & Pilanski, J. M. (2014). Political intolerance, right and left. *Political Psychology*, *35*(6), 841–851. https://doi.org/10.1111/j.1467-9221.2012.00926.x

Dixon, J., Durrheim, K., & Thomae, M. (2017). The principle-implementation gap in attitudes towards racial equality (and how to close it). *Advances in Political Psychology*, *38*(S1), 91–126. https://doi.org/10.1111/pops.12393

Doosje, B., van den Bos, K., Loseman, A., Feddes, A. R., & Mann, L. (2012). "My in-group is superior!": Susceptibility for radical right-wing attitudes and behaviors in Dutch youth. *International Journal of Psychology*, *47*(1), 19–28. https://doi.org/10.1111/j.1750-4716.2012.00099.x

Eschert, S., & Simon, B. (2019). Respect and political disagreement: Can intergroup respect reduce the biased

evaluation of outgroup arguments? *PLoS ONE*, *14*(3), Article 0211556. https://doi.org/10.1371/journal.pone.0211556

Estevan-Reina, L., Bukowski, M., Schäfer, L. F., Potoczek, A., Gurbisz, D., & Christ, O. (2024). *Promoting intergroup tolerance in polarized settings: The role of respect norms.* Manuscript submitted for publication.

Forst, R. (2013). *Toleration in conflict.* Cambridge University Press.

Gibson, J. L. (1989). The structure of attitudinal tolerance in the United States. *British Journal of Political Science*, *19*(4), 562–570. https://doi.org/10.1017/S0007123400005640

Gibson, J. L. (1992). Alternative measures of political tolerance: Must tolerance be "least-liked"? *American Journal of Political Science*, *36*(2), 560–577. https://doi.org/10.2307/2111491

Gibson, J. L., Claassen, C., & Barceló, J. (2020). Putting groups back into the study of political intolerance. In E. Borgida, C. M. Federico, & J. M. Miller (Eds.), *At the forefront of political psychology: Essays in honor of John L. Sullivan* (pp. 55–78). Routledge. https://doi.org/10.4324/9780429351549-5

Haas, I. J., & Cunningham, W. A. (2014). The uncertainty paradox. *Political Psychology*, *35*, 291–302. https://doi.org/10.1111/pops.12035

Herold, M., Janine, J., Otteni, C., & Vorländer, H. (2023). *Polarization in Europe. An analysis of ten European countries.* Mercator Forum Migration and Democracy (MIDEM). https://forummidem.de/wpcontent/uploads/2023/08/TUD_MIDEM_Study_20231_Polarization_in_Europe_.pdf

Hjerm, M., Eger, M. A., Bohman, A., & Connolly, F. F. (2020). A new approach to the study of tolerance: Conceptualizing and measuring acceptance, respect, and appreciation of difference. *Social Indicators Research*, *147*, 897–919. https://doi.org/10.1007/s11205-019-02176-y

Iyengar, S., Lelkes, Y., Levendusky, M., Malhotra, N., & Westwood, S. (2019). The origins and consequences of affective polarization in the United States. *Annual Review of Political Science*, *22*, 129–146. https://doi.org/10.1146/annurev-polisci-051117-073034

Iyengar, S., Sood, G., & Lelkes, Y. (2012). Affect, not ideology: A social identity perspective on polarization. *Public Opinion Quarterly*, *76*(3), 405–431. https://doi.org/10.1016/j.cobeha.2020.07.005

Jost, J. T. (2017). Ideological asymmetries and the essence of political psychology. *Political Psychology*, *38*(2), 167–208. https://doi.org/10.1111/pops.12407

Jost, J. T., Baldassarri, D. S., & Druckman, J. N. (2022). Cognitive–motivational mechanisms of political polarization in social-communicative contexts. *Nature Reviews Psychology*, *1*(10), 560–576. https://doi.org/10.1038/s44159-022-00093-5

Koudenburg, N., & Kashima, Y. (2022). A polarized discourse: Effects of opinion differentiation and structural differentiation on communication. *Personality & Social Psychology Bulletin*, *48*(7), 1068–1086. https://doi.org/10.1177/01461672211030816

Lelkes, Y. (2016). Mass polarization: Manifestations and measurements. *Public Opinion Quarterly*, *80*(S1), 392–410. https://doi.org/10.1093/poq/nfw005

Meleady, R. (2021). "Nudging" intergroup contact: Normative social influences on intergroup contact engagement. *Group Processes & Intergroup Relations*, *24*(7), 1180–1199. https://doi.org/10.1177/13684302211016047

Miller, D. T., & Prentice, D. A. (2016). Changing norms to change behavior. *Annual Review of Psychology*, *67*, 339–361. https://doi.org/10.1146/annurev-psych-010814-015013

Murrar, S., Campbell, M. R., & Brauer, M. (2020). Exposure to peers' pro-diversity attitudes increases inclusion and reduces the achievement gap. *Nature Human Behaviour*, *4*(9), 889–897. https://doi.org/10.1038/s41562-020-0899-5

Neuner, F. G., & Ramirez, M. D. (2023). Evaluating the effect of descriptive norms on political tolerance. *American Politics Research*, *51*(6), 701–714. https://doi.org/10.1177/1532673X231168347

Nolan, J. M. (2021). Social norm interventions as a tool for pro-climate change. *Current Opinion in Psychology*, *42*, 120–125. https://doi.org/10.1016/j.copsyc.2021.06.001

Prentice, D., & Paluck, E. L. (2020). Engineering social change using social norms: Lessons from the study of collective action. *Current Opinion in Psychology*, *35*, 138–142. https://doi.org/10.1016/j.copsyc.2020.06.012

Renger, D., Renger, S., Miché, M., & Simon, B. (2017). A social recognition approach to autonomy: The role of equality-based respect. *Personality and Social Psychology Bulletin*, *43*(4), 479–492. https://doi.org/10.1177/0146167216688212

Renger, D., & Simon, B. (2011). Social recognition as an equal: The role of equality-based respect in group life. *European Journal of Social Psychology*, *41*(4), 501–507. https://doi.org/10.1002/ejsp.814

Riek, B. M., Mania, E. W., & Gaertner, S. L. (2006). Intergroup threat and outgroup attitudes: A meta-analytic review. *Personality and Social Psychology Review*, *10*(4), 336–353. https://doi.org/10.1207/s15327957pspr1004_4

Scanlon, T. M. (2003). The diversity of objections to inequality. In T. M. Scanlon (Ed.), *The difficulty of tolerance: Essays in political philosophy* (pp. 202–218). Cambridge University Press. https://doi.org/10.1017/CBO9780511615153.012

Schäfer, L. F., Tausch, N., Bukowski, M., Jaspers, E., Lubbers, M. J., van Zalk, M., Ciordia, A., Potoczek, A., Estevan-Reina, L., Shani, M., Simons, J.-W., Friehs, M.-T., Gurbisz, D., Middendorf, W., Schäfer, S. J., Ullenboom, J., Graf, S., Hjerm, M., Lavest, C., . . . Christ, O. (2024). *Inclusivity norms counter the negative societal effects of opinion polarization across 12 European countries*. Manuscript submitted for publication. https://osf.io/preprints/psyarxiv/9gh6k_v1

Simon, B. (2023). Taking tolerance seriously: A proposal from a self-categorization perspective on disapproval and respect. *American Psychologist*, *78*(6), 729–742. https://doi.org/10.1037/amp0001166

Simon, B., Eschert, S., Schaefer, C. D., Reininger, K. M., Zitzmann, S., & Smith, H. J. (2019). Disapproved, but tolerated: The role of respect in outgroup tolerance. *Personality and Social Psychology Bulletin*, *45*(3), 406–415. https://doi.org/10.1177/0146167218787810

Simon, B., & Schaefer, C. D. (2016). Tolerance as a function of disapproval and respect: The case of Muslims. *The British Journal of Social Psychology*, *55*(2), 375–383. https://doi.org/10.1111/bjso.12137

Simon, B., & Schaefer, C. D. (2018). Muslims's tolerance towards outgroups: Longitudinal evidence for the role of respect. *The British Journal of Social Psychology*, *57*(1), 240–249. https://doi.org/10.1111/bjso.12213

Skitka, L. J., Bauman, C. W., & Mullen, E. (2004). Political tolerance and coming to psychological closure following the September 11, 2001, terrorist attacks: An integrative approach. *Personality and Social Psychology Bulletin*, *30*(6), 743–756. https://doi.org/10.1177/0146167204263968

Skitka, L. J., Hanson, B. E., Washburn, A. N., & Mueller, A. B. (2013). Exploring the cross-cultural generalizability and scope of morally motivated intolerance. *Social Psychological and Personality Science*, *4*(3), 324–331. https://doi.org/10.1177/1948550612456404

Sullivan, J. L., Piereson, J., & Marcus, G. E. (1979). An alternative conceptualization of political tolerance: Illusory increases 1950s–1970s. *American Political Science Review*, *73*(3), 781–794. https://doi.org/10.2307/1955404

Tankard, M. E., & Paluck, E. L. (2016). Norm perception as a vehicle for social change. *Social Issues and Policy Review*, *10*, 181–211. https://doi.org/10.1111/sipr.12022

Thøgersen, J. (2006). Norms for environmentally responsible behaviour: An extended taxonomy. *Journal of Environmental Psychology*, *26*(4), 247–261. https://doi.org/10.1016/j.jenvp.2006.09.004

Tropp, L. R., O'Brien, T. C., & Migacheva, K. (2014). How peer norms of inclusion and exclusion predict children

interest in cross-ethnic friendships. *Journal of Social Issues*, *70*(1), 151–166. https://doi.org/10.1111/josi.12052

Van Boven, L., Judd, C. M., & Sherman, D. K. (2012). Political polarization projection: Social projection of partisan attitude extremity and attitudinal processes. *Journal of Personality and Social Psychology*, *103*(1), 84–100. https://doi.org/10.1037/a0028145

Velthuis, E., Verkuyten, M., & Smeekes, A. (2021). The different faces of social tolerance: Conceptualizing and measuring respect and coexistence tolerance. *Social Indicators Research*, *158*, 1105–1125. https://doi.org/10.1007/s11205-021-02724-5

Verkuyten, M. (2023). *The social psychology of tolerance*. Routledge. https://doi.org/10.4324/9781003335030

Verkuyten, M., Adelman, L., & Yogeeswaran, K. (2020). The psychology of intolerance: Unpacking diverse understandings of intolerance. *Current Directions in Psychological Science*, *29*(5), 467–472. https://doi.org/10.1177/0963721420924763

Verkuyten, M., & Yogeeswaran, K. (2017). The social psychology of intergroup toleration: A roadmap for theory and research. *Personality and Social Psychology Review*, *21*(1), 72–96. https://doi.org/10.1177/1088868316640974

Verkuyten, M., Yogeeswaran, K., & Adelman, L. (2023). The social psychology of intergroup tolerance and intolerance. *European Review of Social Psychology*, *34*(1), 1–43. https://doi.org/10.1080/10463283.2022.2091326

Westfall, J., Van Boven, L., Chambers, J. R., & Judd, C. M. (2015). Perceiving political polarization in the United States: Party identity strength and attitude extremity exacerbate the perceived partisan divide. *Perspectives on Psychological Science*, *10*(2), 145–158. https://doi.org/10.1177/1745691615569849

Wilson, A. E., Parker, V. A., & Feinberg, M. (2020). Polarization in the contemporary political and media landscape. *Current Opinion in Behavioral Sciences*, *34*, 223–228. https://doi.org/10.1016/j.cobeha.2020.07.005

Zitzmann, S., Loreth, L., Reininger, K. M., & Simon, B. (2022). Does respect foster tolerance? (Re)analyzing and synthesizing data from a large research project using meta-analytic techniques. *Personality & Social Psychology Bulletin*, *48*(6), 823–843. https://doi.org/10.1177/01461672211024422

About the Authors

Lucía Estevan-Reina is a substitute temporary professor in the Department of Social Psychology at the University of Granada. Her main research interest focuses on intergroup relations and social change, especially in the context of cross-group solidarity. Photo: FernUniversität/Thomas Rosenthal ©

Laura Frederica Schäfer is a PhD candidate at the University of Hagen. Her research focuses on intergroup relations, with a particular interest in social norms, intergroup tolerance, and its limits. Photo: FernUniversität/Thomas Rosenthal ©

Wilma Middendorf is a PhD candidate in developmental psychology at Osnabrück University. Her dissertation explores how children develop intergroup attitudes and influence their parents and peers. Photo: FernUniversität/Thomas Rosenthal ©

Marcin Bukowski is a professor at the Jagiellonian University (Institute of Psychology) and a member of the Centre for Social Cognitive Studies in Kraków, Poland. His research interests focus mainly on motivated social cognition and the psychology of intergroup relations. Photo: FernUniversität/Thomas Rosenthal ©

Maarten van Zalk is a full professor of developmental psychology at Osnabrück University who focuses on adolescent development of intergroup relations, social relationships, and antisocial behaviour. Photo: FernUniversität/Thomas Rosenthal ©

Oliver Christ is a full professor at the FernUniversität in Hagen (Faculty of Psychology), Germany. His research interests focus mainly on the social psychology of intergroup relations. Photo: FernUniversität/Thomas Rosenthal ©

ARTICLE

Mind the Gap! Linking Equality-Based Respect Norms with General and Specific Tolerance

Dominika Gurbisz [1,2], **Anna Potoczek** [3], **Marcin Bukowski** [1], **Lucía Estevan-Reina** [4], and **Oliver Christ** [5]

[1] Institute of Psychology, Jagiellonian University, Poland
[2] Doctoral School in the Social Sciences, Jagiellonian University, Poland
[3] Interdisciplinary Centre for Mathematical and Computational Modelling, University of Warsaw, Poland
[4] Department of Social Psychology, University of Granada, Spain
[5] Faculty of Psychology, FernUniversität in Hagen, Germany

Correspondence: Dominika Gurbisz (dominika.gurbisz@doctoral.uj.edu.pl)

Abstract

This research investigates the extent of Poles' tolerance of Jews through the lens of equality-based respect norms and how these norms translate into general and specific tolerance. Additionally, we reexamine the principle-implementation gap, specifically highlighting the potential discrepancy between declared tolerance and actual behavioral intentions. Through our analysis, we explore the distinction between measuring tolerance as an abstract principle and its application in concrete, real-world situations. In three correlational studies ($N = 818$), we examined the relationship between perceived equality-based respect social norms and the declared endorsement of tolerant principles. Studies 1 and 2 sought to establish initial evidence of this association, hypothesizing that individuals are more likely to express tolerance towards Jewish practices when perceiving stronger societal norms emphasizing respect. Results showed that perceived prescriptive equality-based respect within both national and acquaintance ingroup norms was positively linked to tolerant attitudes. Additionally, the findings highlight the complexity of measuring tolerance, revealing a possible overestimation of tolerant attitudes when assessed at an abstract level, compared to when measured through concrete examples.

Keywords

equality-based respect; Jews; polarization; social norms; tolerance

1. Introduction

Fostering tolerance is vital in increasingly diverse, polarized societies (Van Bavel et al., 2024; Vollhardt, 2012). Unlike approaches focused on changing attitudes toward outgroups—which have shown limited success in reducing anti-democratic tendencies (Voelkel et al., 2022)—tolerance does not require fostering positive feelings toward outgroups but rather accepting their rights, beliefs, and practices (Verkuyten, 2010). This is especially true for respect-based tolerance, rooted in respecting others as disapproved equals (Hjerm et al., 2019; Simon, 2023; Velthuis et al., 2021). One way to cultivate such tolerance is through equality-based respect norms. Recent research links these norms to greater tolerance toward disapproved groups (Estevan-Reina et al., 2024a, 2024b; Schäfer et al., 2024).

However, tolerance is often studied broadly (Van Doorn, 2014), focusing on general principles that may differ from assessments in specific contexts (Jackman, 1978; Lawrence, 1976; Verkuyten, 2023). Researchers have increasingly explored whether assessing tolerance through concrete examples provides a more nuanced understanding of attitudes and intentions (Adelman & Verkuyten, 2019; Van Der Noll, 2013; Verkuyten & Slooter, 2007). Building on this, this article presents two studies that examine the potential gap between general tolerance (GT) and specific tolerance (ST), specifically whether equality-based respect norms are linked with GT and ST. Previous findings revealed the link between respect norms and GT (Estevan-Reina et al., 2024a; Schäfer et al., 2024). This research extends these findings by examining whether this relationship also holds for ST. Additionally, we test how the proximity of respect norms—national vs. acquaintance—relates to both GT and ST. Since acquaintance norms may differ from national norms (Gerber et al., 2008), this proximity-based variation highlights the need of examining local vs. national norms in relation to tolerance, specifically towards Jews in Poland.

2. Theoretical Framework

2.1. The Notion of Tolerance

Tolerance is undeniably popular—reflected in politics, pop culture, academic work, and dedicated days (Verkuyten, 2023). This focus is warranted, as tolerance involves continuously weighing reasons to accept disapproved perspectives or practices (Gardner, 1993; Horton, 1996; Verkuyten et al., 2019), fostering coexistence in diverse societies. Tolerance can be understood at two levels: generally, as a principle, and specifically, as applied to concrete cases (Jackman, 1978). For both levels the same conditions apply—one cannot tolerate something to which one is indifferent, neutral, or open to (Cohen, 2004; Gibson, 2006). An additional complexity arises when individuals apply these principles in practice. This challenge is known as the "principle-implementation gap." Dixon et al. (2017) show that people may endorse abstract principles like equality but resist actions to implement them. Lawrence (1976) similarly highlights the gap between abstract commitments to tolerance and their real-world application, especially when core values are challenged. Genuine tolerance is not merely an intellectual exercise—it demands a willingness to tolerate specific, disapproved beliefs and practices. The disapproval–respect model of tolerance suggests that respect serves almost as a prerequisite for tolerance (Simon, 2023; Simon & Schaefer, 2016). However, tolerance may also be possible without respect, rooted in other values such as coexistence and esteem (Forst, 2018). Coexistence-based tolerance is driven by a pragmatic desire for peaceful interaction, without necessarily valuing differences (Forst, 2018). Respect-based tolerance is rooted in a moral acknowledgment

of others' rights and differences (Hjerm et al., 2019). Esteem-based tolerance involves an appreciation of others' beliefs or customs. In line with this, there are scales that focus on the reasons to tolerate (acceptance, respect, and appreciation; Hjerm et al., 2019). Some authors found that respect-based tolerance is a more principled approach to tolerance than pragmatic alternatives (Velthuis et al., 2022). Beyond reasons for tolerance, individuals may selectively accept certain practices of a group while rejecting others based on the values these practices reflect (Gibson & Gouws, 2003).

Measuring only abstract endorsements of tolerance risks overestimating actual societal acceptance. Verkuyten and Slooter (2007) further explored this divide through vignettes, highlighting how political tolerance can differ between values and practices. Although evidence shows tolerance depends on specific beliefs or practices rather than being a universal concept, it is often measured as if it were (Jackman, 1978; Van Doorn, 2014). Our research addresses this gap by examining concrete expressions of tolerance rather than solely abstract principles. We seek to explore the discrepancy between people's declarations of tolerance as general principles vs. specific situations. In the tolerance items, we have also embedded the types of tolerance (coexistence, respect, and appreciation), focusing on them in a combined way to offer a more comprehensive understanding of tolerance. Additionally, this research offers a unique perspective on tolerance toward Jewish practices and beliefs from the viewpoint of Poles, addressing a scarcity of data available on this topic.

2.2. Equality-Based Respect Norms

Individuals are shaped by their environments. The deliberative nature of tolerance suggests that societal norms can shape and modify individual tolerance levels over time (Verkuyten et al., 2021). Building on this, the current study is grounded in two fundamental principles: that social norms influence beliefs and behaviors, and that a relationship exists between social norms and tolerance (Cialdini & Trost, 1998; Neuner & Ramirez, 2023). Norms can influence judgments based on both well-reasoned principles and heuristic understandings (Verkuyten & Yogeeswaran, 2016). This research focuses on the role of equality-based respect norms which emphasize recognizing others as equal participating members in society (Rawls, 2001; Simon & Schaefer, 2016). Respect is crucial for fostering tolerance, as it involves acknowledging others' autonomy, and ability to make independent choices, even in disagreement (Cohen, 2004; King, 1971).

This research also considers the social distance of norms, drawing on self-categorization theory (Tajfel & Turner, 1979), which posits that norms from closer identity groups influence attitudes and behaviors more than distant norms (Gerber et al., 2008; Gerber & Rogers, 2009). This principle also informs the Disapproval-Respect Model of Tolerance, where emphasizing shared identities strengthens the influence of respect norms across groups (Simon, 2023). Studies in Poland, Spain, and Chile confirm that norms from close social circles have a stronger impact on behavioral intentions than broader national norms (Potoczek et al., 2022). This proximity-based variation in normative influence underscores the need to examine how specific social norms shape tolerance in different contexts. The relationship between norms and tolerance is complex; exposure to intolerant descriptive norms, may not lead to more intolerance and may, in some cases, "backfire," increasing tolerance (Neuner & Ramirez, 2023). This study explores how equality-based respect norms shape tolerance, focusing on prescriptive norms, which research suggests have a stronger impact on attitudes than descriptive norms (Schäfer et al., 2024; Smith & Louis, 2008).

3. Current Research

Poland has historically been characterized as a land of religious and cultural coexistence, earning the label "state without stakes" (Tazbir, 2005) for its relative tolerance during periods of widespread persecution in Europe. By the 18th century, Poland was home to the majority of Europe's Jews (Tomaszewski, 1990). However, the Nazi occupation led to the near-total destruction of Jewish life, and post-war tensions, fueled by economic and ideological divisions, gave rise to enduring prejudices (Bilewicz et al., 2012). While recent reports indicate a positive shift in attitudes towards Jews (CBOS, 2021), contemporary research suggests that antisemitism in Poland has evolved into subtler, more structural forms (Bilewicz, 2020). Historically, tolerance research has emphasized abstract principles over specific cases. While people often express general support for tolerance, this tends to weaken when confronted with concrete policies or actions (Jackman, 1978; Lawrence, 1976). Previous work predominantly examined prejudice and discrimination providing a limited view of tolerance towards Jewish beliefs and customs in Poland, particularly how tolerance operates in practice. Though overt intolerance towards Jewish beliefs and practices is rare and typically isolated (ODIHR, 2023), underlying biases persist, often less visible or acknowledged. Many Poles may be unaware of, or unwilling to recognize, these latent prejudices.

To address the gap between declared support for general statements and responses in concrete situations, this research pursues two parallel goals: (a) to elaborate on the current state of knowledge related to measuring tolerance and (b) to explore the relationship between respect norms and different tolerance levels. Specifically, we compare two levels of expressed tolerance towards Jews in Poland: general (GT) and specific (ST). Moreover, we investigate the relationship between national and immediate social surroundings (Study 1: acquaintance; Study 2b: personal environment) equality-based respect norms with both levels of tolerance. An additional goal is developing and validating a new tolerance scale. Study 1 takes an exploratory approach, offering initial insights and testing a preliminary version of the scale, whereas Study 2 adopts a more confirmatory approach to further assess its reliability and validity.

Furthermore, the research explores whether the association between respect norms and tolerance holds when tolerance is measured at a specific, rather than general, level (Estevan-Reina et al., 2024a; Schäfer et al., 2024). In Study 1, we examine Polish attitudes toward Jews, equality-based prescriptive norms, and both GT and ST. In Study 2a, we refine the ST scale changing the contexts to better depict the challenges and issues Jews face, additionally, we evaluate two distinct disapproval measures. Study 2b aims to replicate the main hypotheses tested in Study 1, incorporating disapproval as an exclusion criterion. Specifically, it examines differences between GT and ST, as well as the varying relationships between respect norms at different levels of social proximity and tolerance. The terms used in Studies 1 and 2b, such as "acquaintance norms" and "personal environment norms," were intentional refinements to capture subtle variations in how individuals perceive norms within their immediate social surroundings. The primary distinction is reference group proximity: The studies examine norms derived from individuals' immediate social environment, as opposed to broader national norms, representing a more generalized reference group (Bicchieri et al., 2022). The ST scale is further refined by assessing tolerance for each context individually. Studies 2a and 2b were conducted during periods of the intensification of the Israel-Palestinian conflict (detailed information can be found in the Supplementary Materials). The hypotheses for Study 2b were pre-registered and are available on AsPredicted (https://aspredicted.org/vk4f-7w5b.pdf). We tested the following hypotheses:

H1. On average, participants will score significantly higher on the GT scale than on the ST scale (Studies 1, 2a, and 2b).

H2. The positive correlation between prescriptive acquaintance equality-based respect norms and GT will be stronger than the positive correlation between prescriptive national equality-based respect norms and GT (Studies 1 and 2b).

H3. The positive correlation between prescriptive acquaintance equality-based respect norms and ST will be stronger than the positive correlation between prescriptive national equality-based respect norms and ST (Studies 1 and 2b).

4. Study 1

4.1. Method

4.1.1. Participants

The sample was collected through a snowballing procedure on social media (e.g., Facebook), with 302 initial respondents. Participants who failed attention checks were automatically excluded. After applying exclusion criteria—failing an attention check or practicing Judaism (see Supplementary File, Table 1)—the final sample included 300 participants (for descriptive statistics see Table 1). Participants were informed that the study concerns opinions about the coexistence of social groups.

4.1.2. Measures

Participants completed a feelings thermometer to rate their sentiments toward Jews on a scale from 0 (unfavorable) to 100 (favorable). This measure has been used in Poland (Marmola & Darmoń, 2022), with preliminary evidence suggesting greater reliability in online formats (Liu & Wang, 2015). Equality-based respect norms were assessed using four items (Renger et al., 2017; Schäfer et al., 2024; Simon & Schaefer, 2017), two measuring national norms (e.g., "Most Poles believe that Jews should always be treated as people of equal value"; $\alpha = .86$), while two measured acquaintance norms (e.g., "Most of my acquaintances believe that Jews should always be treated as people of equal value"; $\alpha = .95$). Respondents were given a 7-point Likert scale, from 1 (*strongly disagree*) to 7 (*strongly agree*).

Tolerance was measured using two 9-item scales: GT and ST, each incorporating three types of tolerance (coexistence, respect, appreciation) embedded within the items. Due to a lack of consistent distinction between types (see Supplementary File, Tables 3–9) they were aggregated into their corresponding levels: GT and ST. GT (based on Hjerm et al., 2019; Velthuis et al., 2021) assessed principle-based tolerance, e.g., "The possibility for Jews to live according to their own values, is crucial to ensure harmonious coexistence in society" ($\alpha = .94$). ST measured tolerance in three contexts: "Jews in Poland should be able to receive their own religious education in public schools"; "Jews in Poland should have access to a separate prayer room at work"; "Jews in Poland should be able to have a day off from work and school on Yom Kippur" ($\alpha = .91$). The scales captured both GT and ST with responses on a 7-point Likert scale from 1 (*strongly disagree*) to 7 (*strongly agree*). Details regarding scale development and factor structure are in the Supplementary File.

Table 1. Descriptive statistics for Studies 1, 2a, and 2b.

Variables	Labels	Study 1	Study 2a	Study 2b
Number of participants		300	79	439
Gender	Male	19% ($n = 57$),	17.7% ($n = 14$)	45.6% ($n = 200$)
	Female	79.3% ($n = 238$)	78.5% ($n = 62$)	54.4% ($n = 239$)
	Other	1.7% ($n = 5$)	3.8% ($n = 3$)	0% ($n = 0$)
Age		Not explicitly measured; participants were required to be 18 years or older as part of the inclusion criteria.	34.38 ($SD = 13.98$)	42.12 ($SD = 13.41$)
Disapproval towards Jewish practices/beliefs		0.34 ($SD = 1.16$)	Positive/Negative: 0.67 ($SD = 1.59$) Agreement: 0.18 ($SD = 1.56$)	−2.69 ($SD = 1.46$)
	Type of Scale	1 item used (from −3 to +3).	Two items used (−3 to 3). The first item asks for positive/negative judgment; the second for agreement.	1 item used (from −5 to 5) as an exclusion criterion—only those who scored between −5 and −1 were allowed to continue.
Attitudes towards Jews		77.63 ($SD = 22.69$)	Not measured	33.29 ($SD = 24.28$)

Behavioral intentions were measured using a 5-item scale ($\alpha = .86$) assessing willingness to act in accordance with tolerance, e.g., responding to anti-Semitic comments or signing a petition to help the Jewish community. Items were rated on a 7-point Likert scale (1 = *very unlikely*, 7 = *very likely*) and reflect active forms of tolerance (Adelman et al., 2022).

Disapproval of Jewish practices was measured with a single item from Simon and Schaefer (2016), where participants evaluated Jewish practices/beliefs on a 7-point scale ranging from −3 (*negative*) to +3 (*positive*), with 0 as neutral.

4.2. Results

To test H1, we examined the relationship between ST and GT. As GT violated the normality assumption (Shapiro-Wilk, $p < .05$), a Wilcoxon signed-rank test was conducted. GT ($M = 5.38$, $SD = 1.30$) was significantly higher than ST ($M = 4.28$, $SD = 1.54$), $Z = −12.89$, $p < .001$, with a large effect size ($r = 0.75$) supporting H1 (see Figure 1). Additional analyses across the three types of tolerance (coexistence, respect, and appreciation), revealed consistent patterns (see Supplementary File, page 35).

To test H2 and H3, we investigated the associations between national and acquaintance equality-based respect norms with both GT and ST (see Table 2). Using Soper's (2025) calculator, the correlation between
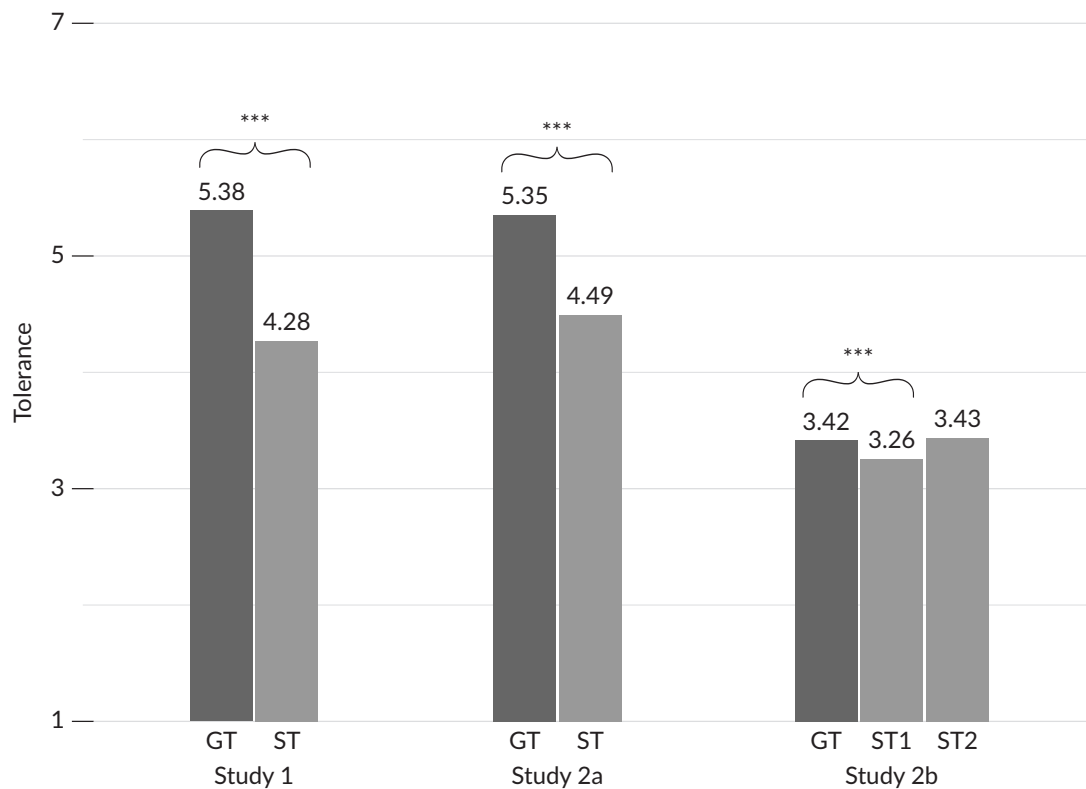
Figure 1. Distribution of general and specific tolerance scores in Studies 1, 2a, and 2b. Notes: General tolerance (GT), specific tolerance (ST), specific tolerance context 1 (ST1), specific tolerance context 2 (ST2).

national respect norms and GT ($r = .26$) was significantly lower than that of acquaintance respect norms and GT ($r = .51$; $z = -3.61$, $p < .001$), supporting H2.

Similarly, national norms ($r = .14$) correlated more weakly with ST than acquaintance norms ($r = .38$; $z = -3.16$, $p = .002$) confirming H3. A similar pattern was observed for behavioral intentions, with national norms ($r = .18$) showing a weaker correlation than acquaintance norms ($r = .449$, $z = -4.31$, $p < .001$). Given the religious influences within the ST contexts, we conducted additional analyses controlling for secularism, which yielded consistent findings (see Supplementary File, Additional Analyses Controlling for Secularism in Study 1).

Table 2. Correlations between equality-based respect norms and social tolerance for Study 1.

Variable	*M*	*SD*	1	2	3	4	5
1. National norms	4.55	1.41					
2. Acquaintance norms	5.79	1.49	.51**				
3. GT	5.38	1.30	.26**	.51**			
4. ST	4.28	1.54	.14*	.38**	.68**		
5. Behavioral intentions	3.95	1.54	.18**	.49**	.57**	.55**	

To further investigate these hypotheses, we conducted exploratory regressions with both equality-based respect norms as simultaneous predictors (see Table 3). Acquaintance respect norms were more predictive of both GT and ST than national norms. This pattern also held for behavioral intentions.

Table 3. Regression analysis of national and acquaintance equality-based respect norms on tolerance in study 1.

Outcome	Predictors	b	SE	t
GT	National	0.01	0.05	0.20
	Acquaintance	0.27***	0.05	5.37
ST	National	−0.08	0.06	−1.32
	Acquaintance	0.22***	0.06	3.56
Behavioral intentions	National	−0.10	0.06	−1.74
	Acquaintance	0.35***	0.06	5.96

Notes: Across this article, controlling for disapproval, education, and attitudes towards Jews; significance levels: *** $p < .001$, ** $p < .01$, * $p < .05$; national equality-based respect norms (National), acquaintance equality-based respect norms (Acquaintance).

To assess whether respect norms relate differently to ST versus GT we conducted Fisher's Z-tests using the Soper calculator. The correlation between national norms and GT ($r = .26$) did not differ significantly from that with ST ($r = .14$), $z = 1.53$, $p = .127$. However, acquaintance respect norms showed a higher correlation with GT ($r = .51$) than with ST ($r = .38$), $z = 1.98$, $p = .047$.

Building on the factor analysis (detailed in the Supplementary File), which confirmed ST as distinct and adding explanatory value beyond GT, we evaluated the significance of this added value by testing its link with behavioral intentions. The regression analysis included the same control variables as in previous regression analyses. Both GT ($b = 0.28$, $t = 3.74$, $p < .001$) and ST ($b = 0.24$, $t = 3.77$, $p < .001$) emerged as significant predictors of behavioral intentions (see summary of analyses in Table 4).

Table 4. Regression analysis of GT and ST on behavioral intentions in Studies 1, 2a, and 2b.

Study	Predictors	b	SE	t
Study 1	GT	0.28***	0.08	3.74
	ST	0.24***	0.06	3.77
Study 2a	GT	0.04	0.07	0.52
	ST	0.26***	0.06	4.53
Study 2b Context 1	GT	0.05	0.03	1.59
	ST	0.15***	0.03	4.92
Study 2b Context 2	GT	0.02	0.04	0.54
	ST	0.17***	0.03	5.46

Notes: Across this paper, controlling for disapproval, education, and attitudes towards Jews, except for Study 2a where attitudes were not measured; significance levels: *** $p < .001$, ** $p < .01$, * $p < .05$.

4.3. Discussion

Study 1 aimed to test the preliminary social tolerance scale and compare mean scores between GT and ST. Our H1 was confirmed: Participants scored higher in GT than in ST. H2 and H3 were also supported: acquaintance norms correlated more strongly with both tolerance levels than national norms. Regression analysis confirmed that acquaintance norms more strongly predict both tolerance levels and behavioral

intentions. Additionally, we have preliminary evidence that respect norms permeate tolerance even when measured at a more specific level.

Exploratory analyses showed that both GT and ST significantly predicted behavioral intentions. However, the sample was skewed, with overwhelmingly positive attitudes toward Jews and only 12% expressing disapproval. The preliminary ST measure included all three contexts within a religious framework, prompting a revision to assess one context at a time. This may have conflated ST with secularism, as items like religious holidays, religious education, and prayer spaces likely overlapped conceptually. Tolerance was measured including the disapproval-approval spectrum as in other tolerance studies (Simon & Schaefer, 2016; Velthuis et al., 2021). The data does not allow exclusion of participants with neutral, indifferent, or positive views about Jewish beliefs/practices, which could inflate tolerance levels as tolerance involves disapproval before acceptance. To address this, Study 2a pretested two disapproval measures, with one serving as an exclusion criterion in Study 2b to assess whether differences between GT and ST persist.

5. Study 2a

In the previous study, disapproval was measured as an exploratory variable but was not incorporated into the analysis due to sample size limitations. However, we consider disapproval a key variable. Traditionally measured using the "least liked" approach (Gibson, 1992), disapproval now extends beyond negative attitudes, informing the bases of weak and strong toleration (Verkuyten & Yogeeswaran, 2016). Current measures assess a group's practices on a good-bad continuum or disapproval of specific practices (e.g., Adelman & Verkuyten, 2019), but separating disapproval of practices from disapproval of the group is challenging. Generalized disapproval signals a deeper rejection of multiple aspects of the group's identity, making it more likely to fuel polarization. Study 2a pre-tests the refined tolerance scale for Study 2b, introducing two disapproval measures, which establish an anchor point for the disapproval exclusion criterion in Study 2b. These items reflect objections to specific beliefs and practices, contrasting with the feeling thermometer. While this distinction was not directly analyzed in Study 2a, it provided a conceptual basis for differentiating between broad group rejection and disapproval of specific practices in Study 2b. The aim of Study 2a is twofold: first, replicate findings on the distribution of GT and ST (H1); and second, refine the tolerance scale by expanding beyond religious contexts. Unlike Study 2b, Study 2a does not examine the relationship between respect norms and tolerance but focuses on pre-testing measures for Study 2b.

5.1. Method

5.1.1. Participants

The sample was collected through a snowballing procedure on social media (e.g., Facebook). A total of 82 respondents participated. After applying the exclusion criteria which included participants practicing Judaism or having a Jewish ethnic background (see Supplementary File, Table 1), 79 participants remained (for descriptive statistics see Supplementary File, Table 11). Participants were informed that the study examined opinions regarding the coexistence of different social groups.

5.1.2. Measures

We assessed two disapproval measures: "Do you consider the beliefs, customs, and behaviors of the Jewish community in Poland to be positive or negative?" (adapted from Simon & Schaefer, 2016) and "Do you agree with the beliefs, customs, and behaviors of the Jewish community in Poland?" Both used a 7-point scale from −3 (*strongly disapprove/negative*) to +3 (*strongly approve/positive*), with 0 as neutral.

Tolerance was measured using two scales—GT and ST—each incorporating distinct types of tolerance: coexistence, respect, and appreciation. The GT scale (9 items, based on Hjerm et al., 2019; Velthuis et al., 2021) captured tolerance for Jews living according to their beliefs and traditions through broad, principle-based terms, e.g., "the possibility for Jews to live according to their own values, is crucial to ensure harmonious coexistence in society" ($\alpha = .94$). The ST scale (9 items, $\alpha = .93$) measured attitudes across three contexts: "In public places, such as the chambers of the Polish parliament, symbols of the Jewish religion, such as a menorah during the celebration of Hanukkah, should also be present" (a menorah is a traditional candelabrum with nine branches); "In all localities where synagogues and Jewish cemeteries are located, local authorities should take care of this heritage in the same way as they would any other sacred site"; "Content related to the history of Polish Jews should be included to a greater extent in Polish textbooks." Responses were recorded on a 7-point Likert scale from 1 (*strongly disagree*) to 7 (*strongly agree*).

Behavioral intentions were assessed using a three-item version ($\alpha = .67$), using a 4-point Likert scale from 1 (*definitely no*) to 4 (*definitely yes*). Items ranged from liking a post promoting Jewish culture to signing a petition regarding the city covering the cost of cleaning Jewish cemeteries.

5.2. Results

To test H1, Shapiro-Wilk tests indicated that GT did not meet the normality assumption ($p < .05$). Wilcoxon signed-rank tests revealed a significant difference between GT ($M = 5.35$, $SD = 1.42$) and ST ($M = 4.49$, $SD = 1.56$), $Z = -6.55$, $p < .001$, with a large effect size ($r = -0.77$; see Figure 1) confirming H1. Additional analyses showed consistent patterns across all three tolerance types (see Supplement File, page 35–36).

As an exploratory analysis, we examined whether ST is more strongly associated with behavioral intentions. In the regression analysis, both ST and GT were included simultaneously as predictors, alongside education and disapproval as controls. The model explained a significant proportion of the variance in behavioral intentions ($\Delta R^2 = .57$, $F(4,74) = 30.08$, $p < .001$). The results indicated that ST was a significant predictor of behavioral intentions ($b = 0.26$, $t = 4.53$, $p < .001$), whereas GT was not ($b = 0.04$, $t = 0.52$, $p = .572$; for detailed results see Table 4).

5.3. Discussion

The main goal of this study was to further refine the ST scale and pretest the disapproval items. As in Study 1, participants reported higher GT than ST. Exploratory regression analysis yielded preliminary findings suggesting that ST, rather than GT, plays a more critical role in predicting behavioral intentions. This result aligns with the theoretical model, emphasizing the distinct and meaningful contribution of ST in understanding behavioral outcomes.

6. Study 2b

This study further examines the hypothesis that participants will score higher on the GT scale than on the ST scale (H1) and employs disapproval as an exclusion criterion. Building on the evidence from Study 1 supporting H2 and H3, we align with the self-categorization theory (Tajfel & Turner, 1979), emphasizing the stronger influence of local norms over national norms on attitudes and behaviors. We extend the concept of local norms from "acquaintance norms" to "personal environment norms" to explore whether the stronger link between acquaintance norms and tolerance observed in Study 1 persists in a slightly broader social context. This provides an opportunity to replicate Schäfer et al. (2024) findings, which showed that equality-based respect norms are associated with higher tolerance, even among those who strongly disapprove of these views. We aim to test these associations between equality-based respect norms and tolerance within a Polish sample and extend this to incorporate the different levels of tolerance. The content of the tolerance measures has been improved, basing them on more relevant issues. The study was pre-registered (https://aspredicted.org/vk4f-7w5b.pdf).

6.1. Methods

6.1.1. Participants

Using G*Power 3.1.9.6 (Faul et al., 2007), we calculated the required sample size to detect a medium effect size ($f = 0.19$) with $\alpha = 0.05$ and power = 0.95, resulting in a target sample of approximately 300 respondents. A total of 445 participants completed the online Polish-language survey, distributed via the Ariadna research panel. Participants who failed the attention checks were automatically prevented from continuing the survey. After applying the pre-registered and post-hoc criteria, six participants were excluded (based on religious affiliation and ethnicity), leaving a final sample of 439 participants.

6.1.2. Measures

We used the same feelings thermometer from Study 1: 0 (indicating cool and unfavorable attitudes) to 100 (indicating warm and favorable attitudes towards Jews).

Disapproval of Jewish practices was assessed using a scale ranging from -5 to $+5$, with the prompt: "I approve/disapprove of the beliefs and customs of the Jewish community in Poland." Participants who indicated neutrality or approval were excluded from the analysis.

We measured equality-based respect norms using four items (Renger et al., 2017; Schäfer et al., 2024; Simon & Schaefer, 2017). Two assessed national norms, e.g., "Most Poles believe that Jews should always be treated as people of equal value" ($\alpha = .88$). Acquaintance norms were replaced with personal environment norms, such as, "Most people in my personal environment believe that Jews should always be treated as people of equal value" ($\alpha = .94$). Respondents indicated their level of agreement on a 7-point Likert scale, ranging from 1 (*strongly disagree*) to 7 (*strongly agree*).

Tolerance was measured using three scales: a 9-item GT scale (adapted from Study 1; $\alpha = .95$) and two 9-item ST scales (18 items total), each assessing a separate context: (a) education ("Content related to the history of

Polish Jews should be included to a greater extent in Polish textbooks" [$\alpha = .97$]) and (b) cultural heritage ("In all localities where synagogues and Jewish cemeteries are located, local authorities should care for this heritage as they would any other sacred site" [$\alpha = .97$]; see factor analyses in the Supplementary File). Each item was rated on a 7-point scale. We also included three exploratory items—one for GT and two for ST—to assess respondents' attitudes toward these contexts without referencing specific types of tolerance. Descriptive statistics and additional analyses are presented in the Supplementary File, Table 10. Behavioral intentions were measured using a revised four-item scale reflecting varying difficulty levels, e.g., "Imagine you come across a post on the Internet promoting Jewish culture. Would you be inclined to like this post?" ($\alpha = .84$). Responses used the same 4-point Likert as in Study 2a. The items were based on the active tolerance measure developed by Adelman et al. (2022). We also measured passive tolerance behavioral intention items; however, low reliability led to their exclusion (see Supplementary File, Table 19.1).

6.2. Results

To test the relationship between ST and GT, we used Wilcoxon signed-rank tests due to GT's non-normal distribution (Shapiro-Wilk test, $p < .05$). Results revealed a significant difference between GT and ST Context 1 which addressed the inclusion of content on the history of Polish Jews in textbooks ($M = 3.42$, $SD = 1.46$ vs. $M = 3.26$, $SD = 1.52$, $Z = -4.85$, $p < .001$, $r = -0.23$). However, no statistically significant difference was found between GT and ST Context 2 which focused on the preservation and care of Jewish heritage sites, such as synagogues and cemeteries, by local authorities ($M = 3.43$, $SD = 1.58$, $Z = -.05$, $p = .962$). These findings partially support H1 (see Figure 1). Additional analyses using types of tolerance revealed a consistent pattern (see Supplementary File, page 36).

To test H2 and H3, we examined the associations between national and personal environment equality-based respect norms with GT and ST (for both contexts; see Table 5). Using the Soper (2025) calculator, we found no significance between personal environment norms ($r = .59$) and national norms ($r = .57$) in predicting GT ($z = -.45$, $p = .656$), thus not supporting H2. Similarly, personal environment norms ($r = .49$) were not significantly more strongly correlated with ST Context 1 than national norms ($r = .49$). The same pattern was observed for ST Context 2 (National Norms: $r = .50$; personal environment norms: $r = .51$; $z = -.20$, $p = .843$), failing to confirm H3.

Table 5. Correlations between equality-based respect norms and social tolerance for Study 2b.

Variable	M	SD	1	2	3	4	5	6
1. National norms	3.44	1.55						
2. Personal environment norms	3.53	1.71	.81**					
3. GT	3.42	1.46	.57**	.59**				
4. ST Context 1	3.26	1.52	.49**	.49**	.87**			
5. ST Context 2	3.43	1.58	.50**	.51**	.89**	.88**		
6. Behavioral intentions	1.72	0.64	.35**	.37**	.59**	.61**	.62**	

To further investigate these hypotheses, we conducted exploratory regressions examining links between equality-based respect norms and tolerance. Both measures of equality-based respect norms were included as simultaneous predictors, while controlling for education, attitudes toward Jews, and disapproval of Jewish

practices. Both national ($b = 0.22$, $p < .001$) and personal environment norms ($b = 0.21$, $p < .001$) significantly predicted GT. ST showed similar patterns: In Context 1, national ($b = 0.21$, $p = .001$) and personal environment norms ($b = 0.13$, $p = .030$) significantly predicted ST; in Context 2, national ($b = 0.23$, $p = .001$) and personal environment norms ($b = 0.16$, $p = .013$) were also significant. Neither national ($b = 0.04$, $p = .126$) nor personal environment norms ($b = 0.03$, $p = .326$) significantly predicted Behavioral Intentions. Table 6 summarizes these results.

Table 6. Regression analysis of national and personal environment equality-based respect norms on tolerance Study 2b.

Outcome	Predictors	b	SE	t
GT	National	0.22***	0.06	3.89
	Personal env.	0.21***	0.05	3.84
ST Context 1	National	0.21**	0.07	3.31
	Personal env.	0.13*	0.06	2.17
ST Context 2	National	0.23**	0.07	3.44
	Personal env.	0.16*	0.06	2.50
Behavioral intentions	National	0.04	0.03	1.53
	Personal env.	0.03	0.03	0.98

To assess whether the respect norm–tolerance link holds across more specific tolerance levels, we conducted Fisher's Z-tests. The correlation between national respect norms and GT ($r = .57$) was not significantly different from ST Context 1 ($r = .49$), $z = 1.65$, $p = .100$; or ST Context 2 ($r = .50$), $z = 1.45$, $p = .147$. However, the correlation between personal environment respect norms, and GT ($r = .59$) was significantly higher than with ST Context 1 ($r = .49$), $z = 2.09$, $p = .037$ but not with ST Context 2 ($r = .51$), $z = 1.70$, $p = .090$.

As an exploratory analysis, we examined whether ST more strongly predicts behavioral intentions. We ran two regressions: one with GT and ST for Context 1 as predictors, and another with GT and ST Context 2, controlling for education and disapproval. In both models, ST significantly predicted behavioral intentions—Context 1: ($b = 0.15$, $t = 4.92$, $p < .001$); Context 2 ($b = 0.17$, $t = 5.46$, $p < .001$)—whereas GT did not (Context 1: $b = 0.05$, $t = 1.59$, $p = .113$; Context 2: $b = 0.02$, $t = 0.54$, $p = .592$; for detailed results see Table 4), replicating Study 1's findings.

6.3. Discussion

The results indicate that participants report higher scores on the GT scale than the ST scale for Context 1, which involved including more content on the history of Polish Jews in textbooks. This finding aligns with our theoretical assumptions, as GT reflects broader, less contextually demanding perspectives. Opposition to such inclusion may reflect more than curricular preference; it could indicate a reluctance to fully acknowledge Jewish history as part of national heritage, thus raising questions about the boundaries of tolerance. This is particularly relevant, as Jewish communities were a significant part of Polish society before WWII. However, this pattern was not observed in Context 2, which involved preserving Jewish heritage sites. One possible explanation for this discrepancy lies in Poland's cultural context. Caring for cemeteries may be less contentious due to the strong tradition of honoring the deceased, especially during All Saints'

Day when people visit cemeteries, decorate graves, and pay their respects. Additionally, the historical context of the Nazi Holocaust likely shapes attitudes toward preserving Jewish heritage sites. In Poland, such efforts may be seen as acknowledging the past, rooted in collective memory. A second possible explanation is that Context 2 might evoke a stronger sense of respect-based tolerance than Context 1. While the context can't be strictly categorized as respect tolerance (or any other type), it may trigger a greater inclination toward tolerance rooted in respect, coexistence, or appreciation. The cultural and historical weight of preserving Jewish heritage sites may be more intuitively linked to values like respecting the past and preserving memory. Although ST Context 2 did not significantly differ from GT, it was more predictive of Behavioral Intentions, suggesting it is more likely to translate into real-world actions.

We replicated Study 1's findings, confirming that Respect Norms are associated with both GT and ST. This is notable given the consistent pattern where participants report higher tolerance for general statements but show decreased tolerance for real-life examples. Despite this, respect norms remain relevant at the specific level, suggesting that they permeate both broad and context-dependent expressions of tolerance. While national respect norms show a stable association with both tolerance levels, personal environment respect norms demonstrate a stronger link with GT than with any specific context. Contrary to Study 1, we found no stronger link between personal environment norms and tolerance. Regression analyses showed that national norms were slightly more predictive of ST than personal environment norms. These findings may reflect limitations of the 'local norm' construct as operationalized in this context, as the high correlation between personal environment norms and national norms suggests that participants viewed them similarly. Though personal environment norms involve immediate surroundings like acquaintance norms, they may lack comparable relational depth.

7. General Discussion

This research explored the relationship between respect norms and tolerance, specifically examining the interplay of GT and ST across contexts. A key finding is that tolerance levels are lower at the specific level compared to general statements, aligning with the "implementation gap" (Dixon et al., 2017), and replicating Jackman's (1978) findings. While individuals may express high tolerance in principle, their acceptance diminishes in real-life scenarios. This pattern was replicated across Studies 1 and 2a, and partially in Study 2b, where the difference was observed only within the educational context.

Despite this overall decline in tolerance at the specific level, the association between respect norms and tolerance remained significant, even when tolerance was measured in more concrete contexts. This finding extends studies on respect norms and tolerance (Estevan-Reina et al., 2024a; Schäfer et al., 2024), as it suggests that equality-based respect norms consistently shape tolerant attitudes, whether assessed broadly or through concrete contexts. While the high correlation between respect norms and tolerant attitudes could reflect projection or motivated reasoning, research on socialization and social learning suggests that external influences, like peer norms, significantly shape tolerance-related attitudes (Simons et al., 2025). This provides preliminary evidence that respect norms serve as a fundamental social mechanism supporting tolerance, even when situational demands challenge individuals' willingness to be tolerant.

An additional central question was whether acquaintance norms are more strongly linked to tolerance than national norms. The findings were mixed. Study 1 supported this hypothesis, showing a stronger association

between acquaintance norms and as well as with behavioral intentions. However, Study 2b, which expanded the scope to broader "personal environment" norms, did not replicate this pattern. The stronger effect of acquaintance norms was not observed, indicating that the impact of social proximity on norm compliance may depend on how narrowly or broadly social groups are defined. This distinction suggests that while respect norms are crucial for promoting tolerance, they may not be sufficient to drive active behavioral intentions. Tolerance reflects a willingness to accept others but does not necessarily translate into emotional engagement or identification with specific practices or groups. Furthermore, these findings contribute to debates in social identity theory (Tajfel, 1982; Terry & Hogg, 1996) and social norm compliance (Bicchieri et al., 2022), emphasizing that closely tied reference groups may exert stronger normative influence on tolerance under specific conditions. However, it is worth noting that significant differences in attitudes toward Jews between Studies 1 and 2b—higher in Study 1 and lower in 2b—could have impacted the relationship between norms and tolerance.

The tolerance scale measuring both GT and ST demonstrated acceptable reliability across studies. Extensive testing, including exploratory and confirmatory factor analyses, and refinement consistently supported a two-dimensional structure (as demonstrated in the Supplementary File), revealing distinct yet correlated factors for GT and ST. A bifactor model provided the best fit, capturing both the overarching construct of tolerance and the unique contributions of each dimension. This model underscored the added explanatory value of ST, particularly in predicting behavioral intentions. Significant differences in participant scores further validated the theoretical distinctiveness of these dimensions, suggesting that while GT provides a general attitudinal foundation, ST offers crucial context-specific insights into expressions of tolerance and their subsequent impact on behavior.

A key contribution of this research is providing deeper insight into the types of norms linked to tolerance and the contextual factors that determine their effectiveness. Specifically, we examine equality-based respect norms as an extension of research linking respect with positive actions toward minority outgroups (Lalljee et al., 2008). In Studies 1 and 2b, we find that equality-based respect norms are associated with ST—a more contextual form of tolerance since it reflects support for concrete actions and policies, rather than just abstract statements. This aligns with previous research demonstrating that respect for outgroups as fellow citizens leads to greater acceptance of concrete practices (Hjerm et al., 2019; Simon et al., 2018). Our research also highlights the importance of norm proximity in shaping tolerance, as people are more likely to follow socially proximate norms without external enforcement, making them more sustainable (Bicchieri et al., 2022). This suggests that norms perceived as distant may be less effective in behavioral interventions. Additionally, understanding the reference group for a norm can help explain why some tolerance norms may backfire (Neuner & Ramirez, 2023).

Tolerance is often criticized for implying hierarchy, casting the tolerated group as deviant or subordinate. However, grounded in equality-based respect, it becomes genuine inclusion rather than permission (Forst, 2018). Unlike coexistence-based interventions focused on attitude change, respect-based norms foster tolerance without requiring such shifts. Prior research has highlighted the shortcomings of coexistence models, which, while promoting peaceful interactions, often fail to address deeper social injustices and power asymmetries (Dixon et al., 2012). For example, in the Israeli-Palestinian context, coexistence initiatives have been criticized for sidelining Arab concerns about structural inequalities, inadvertently reinforcing existing power imbalances (Maoz, 2011). However, the first attempts using equality-based

respect norms as a form of intervention in schools revealed the complexity of the respect norm and tolerance link (Shani et al., 2023). Our findings illustrate that the association is context-dependent, varying based on the reference group proximity, norm specificity, and emotional engagement required for more active behaviors. The link between norms and tolerance is particularly relevant when compared to models like deliberative decision-making (Verkuyten et al., 2022) as evidence suggests in-group norms influence attitude-behavior consistency more reliably than deliberative reasoning alone (Smith & Terry, 2003). This indicates a promising avenue for future research on the role of norms in shaping tolerance, potentially offering a more effective path for fostering social change.

We argue that the general vs. specific distinction is important when studying tolerance, as individuals tend to endorse tolerance in abstract terms rather than in specific situations (Sleijpen et al., 2020). A key contribution of this study is the use of a tolerance scale that aims to address a gap identified in the literature: While much of the existing research focuses on support for broad concepts of tolerance, it often lacks consideration of how tolerance manifests in concrete practices, beliefs, or real-world scenarios (Van Doorn, 2014). This gap is particularly relevant given that research has questioned whether common mechanisms of prejudice reduction, like contact and common identification, consistently translate into tangible behavioral change (Dixon et al., 2012). Our study builds on this perspective by considering whether the discrepancy between professed and practiced tolerance may, in part, stem from the way tolerance is measured. The differences between GT and ST levels, along with the stronger predictive power of ST for behavioral intentions, suggest that general measures may overestimate individuals' levels of tolerance. This underscores the need for more precise tools to capture the complexities of tolerance in practice, particularly in evaluating the effectiveness of tolerance interventions.

8. Limitations and Future Directions

This research is not without its limitations. First, while the studies were conducted before and during the intensification of the Israeli-Palestinian conflict, and importantly, not during periods of heightened attention to this conflict within Poland, we cannot fully rule out the potential influence of such emotionally charged global events on Polish views and attitudes. The data collected does not assess the extent of this influence. Second, although the tolerance scale underwent rigorous refinement and demonstrated high reliability (see Supplementary File), formal validation is still pending. Additionally, the study is subject to self-selection bias, as participants with stronger interest or background knowledge in the topics may have been more likely to participate, particularly in Study 1 which was circulated among internet users. Moreover, the online format excludes individuals without regular internet access and tends to attract a younger demographic. While the online nature offers broad accessibility, it may not fully capture behavioral intentions shaped by offline social networks or face-to-face interactions. Another important consideration is the use of different norm operationalizations across the two studies. Study 1 assessed acquaintance norms, whereas Study 2b focused on personal environment norms. While this distinction limits direct comparability within a single study, it highlights the need to carefully consider how social norms are defined and measured. This research shows that framing norms as acquaintance versus personal environment yields distinct insights, potentially reflecting differences in their strength or nature. In Studies 1 and 2a, several ST items contained embedded religious components, which may have influenced participants' responses in ways not fully accounted for. Although secularism was measured in Study 1 and included as a control variable in supplementary analyses, this approach does not fully eliminate the potential confounding effects of religiosity or secular orientation. As such, the contextual limitations of these measures should be carefully considered in future research.

Moreover, the behavioral intention scale employed in the studies reflects a relatively active form of tolerance, conceptually aligning more closely with measures of collective action.

To preserve the distinctiveness of active tolerance as a construct, future research should refine the measure to more clearly delineate it from collective action tendencies, improving construct validity and enabling a clearer assessment of the tolerance principle implementation gap. Further validation of the tolerance scale is essential to strengthen its psychometric properties. Finally, conducting this research in diverse cultural settings presents a valuable direction for assessing the broader applicability of the findings and deepening our understanding of how social norms influence tolerance and behavioral intentions across different environments.

Conflict of Interests
The authors declare no conflict of interests.

Supplementary Material
Supplementary material for this article is available online in the format provided by the author (unedited).

References
Adelman, L., & Verkuyten, M. (2019). Prejudice and the acceptance of Muslim minority practices. *Social Psychology*, *51*(1), 1–16. https://doi.org/10.1027/1864-9335/a000380

Adelman, L., Verkuyten, M., & Yogeeswaran, K. (2022). Distinguishing active and passive outgroup tolerance: Understanding its prevalence and the role of moral concern. *Political Psychology*, *43*(4), 731–750. https://doi.org/10.1111/pops.12790

Bicchieri, C., Dimant, E., Gächter, S., & Nosenzo, D. (2022). Social proximity and the erosion of norm compliance. *Games and Economic Behavior*, *132*, 59–72. https://doi.org/10.1016/j.geb.2021.11.012

Bilewicz, M. (2020). Moving beyond the past: The role of historical closure in conflict resolution. *Social Psychological Bulletin*, *14*(4). https://doi.org/10.32872/spb.v14i4.2437

Bilewicz, M., Winiewski, M., & Radzik, Z. (2012). Antisemitism in Poland: Psychological, religious, and historical aspects. *Journal for the Study of Antisemitism*, *4*(2), 423–443.

CBOS. (2021). *Stosunek do innych narodów*.

Cialdini, R. B., & Trost, M. R. (1998). Social influence: Social norms, conformity and compliance. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *The Handbook of Social Psychology* (4th ed., pp. 151–192). McGraw-Hill.

Cohen, A. J. (2004). What toleration is. *Ethics*, *115*(1), 68–95. https://doi.org/10.1086/421982

Dixon, J., Durrheim, K., & Thomae, M. (2017). The Principle-Implementation gap in attitudes towards racial equality (and how to close it). *Political Psychology*, *38*(S1), 91–126. https://doi.org/10.1111/pops.12393

Dixon, J., Levine, M., Reicher, S., & Durrheim, K. (2012). Beyond prejudice: Are negative evaluations the problem and is getting us to like one another more the solution? *Behavioral and Brain Sciences*, *35*(6), 411–425. https://doi.org/10.1017/s0140525x11002214

Estevan-Reina, L., Bukowski, M., Schäfer, L., Potoczek, A., Gurbisz, D., & Christ, O. (2024a). *Promoting intergroup tolerance in polarized settings: The role of respect norms*. Manuscript submitted for publication.

Estevan-Reina, L., Schäfer, L. F., Middendorf, W., Bukowski, M., van Zalk, M., & Christ, O. (2024b). *Respect tolerance in polarized contexts: Strength of the association between personal respect norms and tolerance depending on extremity of opinion, strength of disapproval and perceived threat*. Manuscript submitted for publication.

Faul, F., Erdfelder, E., Lang, A., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. https://doi.org/10.3758/bf03193146

Forst, R. (2018). *Toleration, justice and reason*. Manchester University Press. https://doi.org/10.7765/9781526137708.00010

Gardner, P. (1993). *Tolerance and education*. In J. Horton (Ed.), *Liberalism, multiculturalism and toleration* (pp. 83–103). Palgrave Macmillan. https://doi.org/10.1007/978-1-349-22887-4_6

Gerber, A. S., Green, D. P., & Larimer, C. W. (2008). Social pressure and voter turnout: Evidence from a large-scale field experiment. *American Political Science Review*, *102*(1), 33–48. https://doi.org/10.1017/s000305540808009x

Gerber, A. S., & Rogers, T. (2009). Descriptive social norms and motivation to vote: Everybody's voting and so should you. *The Journal of Politics*, *71*(1), 178–191. https://doi.org/10.1017/s0022381608090117

Gibson, J. L. (1992). Alternative measures of political tolerance: Must tolerance be "least-liked"? *American Journal of Political Science*, *36*(2), Article 560. https://doi.org/10.2307/2111491

Gibson, J. L. (2006). Enigmas of intolerance: Fifty years after Stouffer's communism, conformity, and civil liberties. *Perspectives on Politics*, *4*(1). https://doi.org/10.1017/s153759270606004x

Gibson, J. L., & Gouws, A. (2003). *Overcoming intolerance in South Africa: Experiments in democratic persuasion*. Cambridge University Press. https://doi.org/10.1017/cbo9780511550331

Hjerm, M., Eger, M. A., Bohman, A., & Connolly, F. F. (2019). A new approach to the study of tolerance: Conceptualizing and measuring acceptance, respect, and appreciation of difference. *Social Indicators Research*, *147*(3), 897–919. https://doi.org/10.1007/s11205-019-02176-y

Horton, J. (1996). Toleration as a virtue. In D. Heyd (Ed.), *Toleration: An elusive virtue* (pp. 28–43). Princeton University Press. https://doi.org/10.1515/9781400822010.28

Jackman, M. R. (1978). General and applied tolerance: Does education increase commitment to racial integration? *American Journal of Political Science*, *22*(2), Article 302. https://doi.org/10.2307/2110618

King, P. (1971). The problem of tolerance. *Government and Opposition*, *6*(2), 172–207. https://www.jstor.org/stable/44481942

Lalljee, M., Tam, T., Hewstone, M., Laham, S., & Lee, J. (2008). Unconditional respect for persons and the prediction of intergroup action tendencies. *European Journal of Social Psychology*, *39*(5), 666–683. https://doi.org/10.1002/ejsp.564

Lawrence, D. G. (1976). Procedural norms and tolerance: A reassessment. *American Political Science Review*, *70*(1), 80–100. https://doi.org/10.2307/1960325

Liu, M., & Wang, Y. (2015). Data collection mode effect on feeling thermometer questions: A comparison of face-to-face and Web surveys. *Computers in Human Behavior*, *48*, 212–218. https://doi.org/10.1016/j.chb.2015.01.057

Maoz, I. (2011). Does contact work in protracted asymmetrical conflict? Appraising 20 years of reconciliation-aimed encounters between Israeli Jews and Palestinians. *Journal of Peace Research*, *48*(1), 115–125. https://doi.org/10.1177/0022343310389506

Marmola, M., & Darmoń, K. (2022). The political image of left-wing leaders in Poland: How young voters evaluate the politicians from different "political generations." *Political Preferences*, *30*(1), 5–18. https://doi.org/10.31261/polpre.2022.1.5-18

Neuner, F. G., & Ramirez, M. D. (2023). Evaluating the effect of descriptive norms on political tolerance. *American Politics Research*, *51*(6), 701–714. https://doi.org/10.1177/1532673x231168347

ODIHR. (2023). *Poland hate crime report 2023*. https://hatecrime.osce.org/reporting/poland/2023

Potoczek, A., Bukowski, M., De Lemus, S., Jiménez-Moya, G., Rodríguez-López, Á., & Jasko, K. (2022). Walk this way: Ingroup norms determine voting intentions for those who lack sociopolitical control. *Personality and Social Psychology Bulletin*, *49*(5), 692–708. https://doi.org/10.1177/01461672211070070

Rawls, J. (2001). The second part of ideal theory. In J. Rawls (Ed.), *The law of peoples: With "the idea of public reason revisited"* (pp. 59–88). Harvard University Press. https://doi.org/10.2307/j.ctv1pncngc.6

Renger, D., Renger, S., Miché, M., & Simon, B. (2017). A social recognition approach to autonomy. *Personality and Social Psychology Bulletin*, *43*(4), 479–492. https://doi.org/10.1177/0146167216688212

Schäfer, L. F., Tausch, N., Bukowski, M., Jaspers, E., Lubbers, M. J., Van Zalk, M., Ciordia, A., Potoczek, A., Estevan-Reina, L., Shani, M., Simons, J., Friehs, M., Gurbisz, D., Middendorf, W. M., Schäfer, S. J., Ullenboom, J., Graf, S., Hjerm, M., Lavest, C., . . . Christ, O. (2024). *Inclusivity norms counter the negative societal effects of opinion polarization across 12 European countries*. Manuscript submitted for publication. https://doi.org/10.31234/osf.io/9gh6k

Shani, M., De Lede, S., Richters, S., Kleuker, M., Middendorf, W., Liedtke, J., Witolla, S., & Van Zalk, M. (2023). A social network intervention to improve adolescents' intergroup tolerance via norms of equality-based respect: The "together for tolerance" feasibility study. *International Journal of Developmental Science*, *17*(1/3), 93–110. https://doi.org/10.3233/dev-230342

Simon, B. (2023). Taking tolerance seriously: A proposal from a self-categorization perspective on disapproval and respect. *American Psychologist*, *78*(6), 729–742. https://doi.org/10.1037/amp0001166

Simon, B., Eschert, S., Schaefer, C. D., Reininger, K. M., Zitzmann, S., & Smith, H. J. (2018). Disapproved, but tolerated: The role of respect in outgroup tolerance. *Personality and Social Psychology Bulletin*, *45*(3), 406–415. https://doi.org/10.1177/0146167218787810

Simon, B., & Schaefer, C. D. (2016). Tolerance as a function of disapproval and respect: The case of Muslims. *British Journal of Social Psychology*, *55*(2), 375–383. https://doi.org/10.1111/bjso.12137

Simon, B., & Schaefer, C. D. (2017). Muslims' tolerance towards outgroups: Longitudinal evidence for the role of respect. *British Journal of Social Psychology*, *57*(1), 240–249. https://doi.org/10.1111/bjso.12213

Simons, J. W., Jaspers, E., & Van Tubergen, F. (2025). Socialisation and ethnic majorities' attitudes towards ethnic minorities: A systematic review and meta-analysis of correlational evidence. *Journal of Ethnic and Migration Studies*. Advance online publication. https://doi.org/10.1080/1369183x.2025.2472818

Sleijpen, S., Verkuyten, M., & Adelman, L. (2020). Accepting Muslim minority practices: A case of discriminatory or normative intolerance? *Journal of Community & Applied Social Psychology*, *30*(4), 405–418. https://doi.org/10.1002/casp.2450

Smith, J. R., & Louis, W. R. (2008). Do as we say and as we do: The interplay of descriptive and injunctive group norms in the attitude–behaviour relationship. *British Journal of Social Psychology*, *47*(4), 647–666. https://doi.org/10.1348/014466607x269748

Smith, J. R., & Terry, D. J. (2003). Attitude-behaviour consistency: The role of group norms, attitude

accessibility, and mode of behavioural decision-making. *European Journal of Social Psychology*, *33*(5), 591–608. https://doi.org/10.1002/ejsp.172

Soper, D. S. (2025). *Significance of the difference between two correlations calculator* [Software]. https://www.danielsoper.com/statcalc

Tajfel, H. (1982). Social psychology of intergroup relations. *Annual Review of Psychology*, *33*(1), 1–39. https://doi.org/10.1146/annurev.ps.33.020182.000245

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). Brooks/Cole.

Tazbir, J. T. (2005). Marzenia o "Nowej Polsce." *Nauka*, *1*, 25–34.

Terry, D. J., & Hogg, M. A. (1996). Group norms and the attitude-behavior relationship: A role for group identification. *Personality and Social Psychology Bulletin*, *22*(8), 776–793. https://doi.org/10.1177/0146167296228002

Tomaszewski, J. (1990). *Zarys dziejów Żydów w Polsce w latach 1918–1939*.

Van Bavel, J. J., Gadarian, S. K., Knowles, E., & Ruggeri, K. (2024). Political polarization and health. *Nature Medicine*. Advance online publication. https://doi.org/10.1038/s41591-024-03307-w

Van Der Noll, J. (2013). Religious toleration of Muslims in the German Public Sphere. *International Journal of Intercultural Relations*, *38*, 60–74. https://doi.org/10.1016/j.ijintrel.2013.01.001

Van Doorn, M. (2014). The nature of tolerance and the social circumstances in which it emerges. *Current Sociology*, *62*(6), 905–927. https://doi.org/10.1177/0011392114537281

Velthuis, E., Van Der Noll, J., & Verkuyten, M. (2022). Tolerance of Muslim minority identity enactment: The roles of social context, type of action and cultural continuity concern. *Journal of Community & Applied Social Psychology*, *32*(5), 860–871. https://doi.org/10.1002/casp.2605

Velthuis, E., Verkuyten, M., & Smeekes, A. (2021). The different faces of social tolerance: Conceptualizing and measuring respect and coexistence tolerance. *Social Indicators Research*, *158*(3), 1105–1125. https://doi.org/10.1007/s11205-021-02724-5

Verkuyten, M. (2010). Multiculturalism and tolerance: An intergroup perspective. In R. J. C. Crisp (Ed.), *The psychology of social and cultural diversity* (pp. 147–170). Wiley Blackwell. https://doi.org/10.1002/9781444325447.ch7

Verkuyten, M. (2023). *The social psychology of tolerance*. Routledge. https://doi.org/10.4324/9781003335030

Verkuyten, M., Adelman, L., & Yogeeswaran, K. (2021). Tolerance as forbearance: Overcoming intuitive versus deliberative objections to cultural, religious, and ideological differences. *Psychological Review*, *129*(2), 368–387. https://doi.org/10.1037/rev0000228

Verkuyten, M., Schlette, A., Adelman, L., & Yogeeswaran, K. (2022). Deliberative thinking increases tolerance of minority group practices: Testing a dual-process model of tolerance. *Journal of Experimental Psychology Applied*, *29*(2), 414–424. https://doi.org/10.1037/xap0000429

Verkuyten, M., & Slooter, L. (2007). Tolerance of Muslim beliefs and practices: Age related differences and context effects. *International Journal of Behavioral Development*, *31*(5), 467–477. https://doi.org/10.1177/0165025407081480

Verkuyten, M., & Yogeeswaran, K. (2016). The social psychology of intergroup toleration. *Personality and Social Psychology Review*, *21*(1), 72–96. https://doi.org/10.1177/1088868316640974

Verkuyten, M., Yogeeswaran, K., & Adelman, L. (2019). Toleration and prejudice-reduction: Two ways of improving intergroup relations. *European Journal of Social Psychology*, *50*(2), 239–255. https://doi.org/10.1002/ejsp.2624

Voelkel, J. G., Chu, J., Stagnaro, M. N., Mernyk, J. S., Redekopp, C., Pink, S. L., Druckman, J. N., Rand, D. G., &

Willer, R. (2022). Interventions reducing affective polarization do not necessarily improve anti-democratic attitudes. *Nature Human Behaviour*, *7*(1), 55–64. https://doi.org/10.1038/s41562-022-01466-9

Vollhardt, J. R. (2012). Tolerance for diverse groups. In *The Encyclopedia of Peace Psychology* (pp. 1117–1121). Wiley-Blackwell. https://doi.org/10.1002/9780470672532.wbepp279

About the Authors

Dominika Gurbisz is a PhD student at the Doctoral School in the Social Sciences at Jagiellonian University, affiliated with the Institute of Psychology. Her research focuses on the interplay between intergroup toleration and social norms.

Anna Potoczek is a social psychologist. In her scientific work, she conducts research regarding susceptibility to social norms, reactions to crises, and political psychology. Her interests also include good research practices as part of the Open Science movement and the dissemination of scientific knowledge.

Marcin Bukowski is a professor at the Jagiellonian University (Institute of Psychology) and a member of the Centre for Social Cognitive Studies in Kraków, Poland. His research interests focus mainly on motivated social cognition and the psychology of intergroup relations.

Lucía Estevan-Reina is a substitute temporary professor in the Department of Social Psychology at the University of Granada. Her main research interest focuses on intergroup relations and social change, especially in the context of cross-group solidarity.

Oliver Christ is a full professor at the FernUniversität in Hagen (Faculty of Psychology), Germany. His research interests focus mainly on the social psychology of intergroup relations.

cogitatio

ARTICLE

Living Up to Your Own Standards? Patterns of Civic Norms and Volunteering in Germany

Kathrin Ackermann [1], Jonathan Mylius [2], Annette Haussmann [3], and Stefanie Wiloth [4]

[1] Department for Social Sciences, University of Siegen, Germany
[2] Faculty of Theology, Humboldt University of Berlin, Germany
[3] Faculty of Theology, Heidelberg University, Germany
[4] Institute for Ageing Research, Eastern Switzerland University of Applied Sciences, Switzerland

Correspondence: Kathrin Ackermann (kathrin.ackermann@uni-siegen.de)

Abstract

One of the most significant challenges facing contemporary societies is the increasing polarization of social and political ideologies. Against this backdrop, it is highly important to understand the foundations of social cohesion in order to effectively address this trend. One of the fundamental pillars of social cohesion is volunteering, which entails contributing to a collective good through unpaid work in an organization or association. While a substantial body of research has been dedicated to examining the socio-structural and sociodemographic correlates of volunteering, the relationship between norms and volunteering behavior has received comparatively less attention. In this study, we employ the concept of citizenship norms to empirically explore the patterns of civic norms, particularly norms of solidarity and norms of participation, and volunteering. Furthermore, we investigate how these patterns differ across societal groups, including age, gender, education, and religion. Our quantitative analysis is based on data from a population survey in Germany, with approximately 1,800 respondents. We find that civic norms relate to volunteering, with participation norms showing a stronger link than solidarity norms. These relationships are not moderated by moral and socio-structural factors but remain consistent across different societal groups.

Keywords

civic norms; social norms; social capital; solidarity; volunteering

---

1. Introduction

In contemporary Western societies, major crises, such as the Covid-19 pandemic, the Russian war on Ukraine, and the climate crisis, seem to act as catalysts for deep-seated problems. Challenges like perceived social divides, increasing individualization, the impact of social media on trust in political institutions, and the rise of populist parties put both social and political stability to the test (Diamond & Skrzypek, 2024). The most recent examples of these developments are the political actions of the Trump administration in the USA as well as the results of the German federal election of 2025, where the radical right Alternative für Deutschland (AfD) became the second-strongest force in parliament. Against this backdrop, political pundits remind us of what philosophers but also empirical social scientists have repeated over and over since the 19th century: Social cohesion is key to achieving societal and democratic stability (Fonseca et al., 2019). Alexis de Tocqueville identified, in *Democracy in America*, civil society as a space where political processes were negotiated within small entities of society (de Tocqueville, 1835/2000). Walzer (1996) later termed this discovery "political society." According to Walzer, democracies are stronger when rooted in societies with high levels of social cohesion. Social cohesion not only fosters democratic resilience but also enhances economic development (Knack & Keefer, 1997). This conclusion was also prominently shared by Robert D. Putnam in his study of Italian regions (Putnam et al., 2004). Societies with strong social bonds are more resilient, economically stronger, fairer, and better equipped to promote democracy (Dekker & Halman, 2003; Durkheim, 1893/2019; Musick & Wilson, 2008; Putnam, 1995, 2001; Stolle, 2003; Van der Meer & Van Ingen, 2009; Wilkinson & Pickett, 2010; Wilson, 2000).

One of the key aspects of social cohesion is voluntary engagement. According to Wilson (2000, p. 215), "volunteering means any activity in which time is given freely to benefit another person, group, or organization." It is well-understood how sociodemographic and socio-structural factors are linked to volunteering (Wilson, 2000, 2012). Yet, the relationship between norms and volunteering remains fairly understudied. This is surprising given psychological research showing that norms are related to actual behavior (e.g., Ajzen, 1991; Bicchieri, 2006, 2016; Cialdini et al., 1991). Drawing on Dalton's (2008) concept of norms of good citizenship, we aim to help close this research gap and study how these civic norms relate to volunteering. Civic norms are supposed to capture an individual's notion of what it means to be a "good citizen," distinguishing among several norm dimensions: participation, solidarity, autonomy, and social order.

We argue that norms of participation and solidarity, in particular, are positively related to volunteering. Citizens who regard participation and solidarity as integral to their civic duties are more likely to volunteer in order to live up to their own standards. In additional and rather exploratory analyses, we study whether moral and socio-structural factors, like age, gender, education, and religiosity, amplify these relationships. Investigating the interplay between individual characteristics, norms, and voluntary engagement can shed light on how social cohesion might be strengthened. Against the backdrop of the current crises affecting Western democracies, such insights may guide efforts to reinforce social bonds and promote long-term democratic resilience.

To study the link between civic norms and volunteering empirically, we use data from the EXPSOLIDARITY project, a project on values, motives, and practices in social volunteering. In the course of this project, a representative population survey was conducted in two waves between December 2022 and January 2023 ($N \approx 1,800$). Formal volunteering serves as the dependent variable in our analyses and is correlated with

participation norms and solidarity norms taken from Dalton's (2008, 2021) concept of good citizenship. Additionally, moral and socio-structural factors (age, gender, education, religiosity) are included as control and moderating variables.

The empirical analysis demonstrates a positive relationship between civic norms and volunteering, with participation norms showing a stronger link than solidarity norms. A one-point increase in the respective indices is associated with a 10 and 6 percentage point rise in the likelihood of volunteering. However, religiosity and education appear to have an even stronger connection to volunteering than civic norms, while age is not significantly related, and men are more likely to volunteer than women. Further analyses investigate whether the link between civic norms and volunteering varies across societal groups, considering moral and socio-structural factors as moderators. The results show no significant interaction effects, indicating that while the distribution of civic norms differs across groups, their relationship with volunteering remains consistent.

2. The Link Between Civic Norms and Volunteering

Wilson (2000, p. 215) defines volunteering as "any activity in which time is given freely to benefit another person, group, or organization." Volunteering can further be classified into formal and informal types (Ackermann, 2019; Henriksen et al., 2008; Lee & Brudney, 2012; Wang et al., 2017). Formal volunteering involves participation within structured organizations, such as sports clubs or church youth groups, whereas informal volunteering occurs outside of institutionalized contexts. For example, assisting an elderly neighbor is considered informal volunteering. This article focuses on formal volunteering, which typically demands significantly greater time, effort, and commitment from participants compared to informal volunteering (Ackermann, 2019).

Social norms are defined as "the rules and standards that are understood by members of a group, and that guide and/or constrain social behavior without the force of laws" (Cialdini & Trost, 1998, p. 152). For this study, we adopt a specific conceptualization of norms derived from political science research. The concept of citizenship or civic norms has been profoundly influenced by Dalton (2008, p. 78), who defines them as "a set of norms of what people think people should do as good citizens." Dalton's conceptions of norms do not exactly follow the psychological distinction between descriptive and injunctive norms. Yet, the conception is closely related to injunctive norms capturing what individuals think is expected from them in order to be perceived as a "good citizen."

Following this conceptualization, four categories of civic norms can be identified: participation, solidarity, autonomy, and social order (see Table 1). Citizen participation in politics and society is fundamental to democratic systems. It encompasses institutionalized, non-institutionalized, and social forms of engagement. Political participation, in particular, serves as a key source of legitimacy in democratic regimes. Beyond that, societies are expected to benefit from high levels of social participation, as it fosters social cohesion. Solidarity represents the horizontal dimension of civic norms, referring to the ethical and moral responsibility toward fellow citizens. This sense of social responsibility manifests in caring for others and supporting those in need (Zmerli, 2010). It is closely linked to the concept of social citizenship, which emphasizes collective solidarity (Dalton, 2008; Denters et al., 2007). Autonomy pertains to the self-perception of citizens as independent, self-governing individuals. Autonomous citizens actively seek information, critically monitor

government actions, and try to understand the arguments of others in discussions (Dalton, 2008; van Deth, 2009). Finally, social order constitutes the fourth category of civic norms. It reflects the extent to which citizens accept the rule of law and demonstrate a willingness to adhere to legal regulations, such as tax compliance.

Table 1. Categories of civic norms.

Categories	Norms
Participation	Always vote in elections
	Be active in social or political organizations
	Choose products for political, ethical, or environmental reasons
Solidarity	Support people in your own country who are worse off than yourself
	Help people in the rest of the world who are worse off than yourself
Autonomy	Try to understand the reasoning of people with other opinions
	Keep watch on the actions of government
Social Order	Always obey laws and regulations
	Never try to evade taxes

Source: Based on Dalton (2021, p. 27).

We argue that there is a link between two categories of Dalton's (2021) conception of civic norms, namely norms of participation and norms of solidarity, and volunteering. Norms can be perceived as guidelines for behavior that exhibit their effect on behavior through mechanisms like socialization, social learning, and social image and self-image. This theoretical argument is grounded in psychological theories and research. As a fundamental theoretical framework, Ajzen's (1991) theory of planned behavior is particularly useful, as it posits that behavioral intentions, and consequently actions, are shaped in part by subjective norms. Subjective norms refer to a combination of an individual's own normative expectations and their perception of the expectations of significant others regarding a particular behavior.

The relationship between norms and behavior can be theorized in various ways. We draw on Gross and Vostroknutov (2022) to explore potential mechanisms. First, one mechanism for the action-guiding effect of norms is the internalization of norms through socialization. As individuals undergo socialization, they internalize the norms of the dominant social groups surrounding them. Consequently, norm violations elicit negative emotional responses, such as guilt or shame (Giguère et al., 2014). This imprinting process begins in childhood and is reinforced through observations of rewards and punishments associated with adherence to or deviation from social norms (Gavrilets & Richerson, 2017; Horne, 2003).

Second, social learning processes are considered a key mechanism for the relationship between norms and behavior. It is argued that social norms derive their effectiveness from observation, imitation, and adaptation to the behavior of others (Bicchieri & Xiao, 2009; Keizer et al., 2008). Individuals are influenced by the actions of those around them and are more likely to follow norms when they perceive that others adhere to them as well (Bicchieri, 2006; Tremewan & Vostroknutov, 2021). According to Bicchieri (2006), norm adherence depends on three interrelated expectations: Individuals follow norms when they observe others following a norm (empirical expectations), when they have the impression that others expect that they themselves should follow this norm (normative expectations), and when norm violations can be noticed and

sanctioned (conditional expectations). Thus, the mere existence of a norm does not necessarily lead to compliance. Instead, widespread adherence often requires a critical mass of supporters for the norm to become firmly established (Lindström et al., 2018). Conversely, observed norm violations can trigger a downward spiral, in which compliance decreases significantly over time (Bicchieri & Xiao, 2009; Engel & Kurschilgen, 2020).

Third, social image and self-image are considered potential mechanisms influencing norm adherence. Individuals tend to follow norms to be perceived as fair, honest, and trustworthy by those around them, reinforcing their social image (Bursztyn & Jensen, 2017). The desire to be seen positively by others is regarded as a fundamental human motivation (Grimalda et al., 2016). At the same time, people strive to maintain a morally upright self-image, which motivates them to adhere to norms even in the absence of external observation.

To summarize, the link between norms and actual behavior is argued to be largely driven by moral and social obligations that are mainly shaped through processes of socialization, social learning, and self-consistency. Building on these theoretical perspectives, we argue that civic norms are linked to volunteering. Specifically, we assume that participation norms and solidarity norms play a crucial role in motivating individuals to engage in voluntary activities. Volunteering means playing an active role in society. This should be particularly appealing to individuals who hold strong norms of participation. Moreover, volunteering often means contributing to an improvement of living conditions for others, either in their immediate surroundings or in well-off places. These considerations of charity should be particularly relevant for individuals who hold strong norms of solidarity. Thus, we formulate the following two hypotheses that will be tested empirically in the following:

> H1: A person, who strongly supports norms of participation, will be more likely to volunteer.

> H2: A person, who strongly supports norms of solidarity, will be more likely to volunteer.

3. Data, Measurement, and Methods

3.1. Data

To examine the relationship between civic norms and volunteering, we draw on data from the project Experiencing Solidarity: Values, Motives, and Practices in Caring Communities and Social Volunteering (EXPSOLIDARITY). This representative population survey was conducted in two waves between December 2022 and January 2023 using the online access panel of Bilendi & respondi. The primary objective of the survey was to gather data on social cohesion, volunteering, civic norms, motives, and political attitudes in Germany. For this study, we utilize both waves of the panel, resulting in a final sample size of approximately 1,800 respondents. In the initial sample (first wave of the panel), quotas of age, gender, and education have been applied to ensure that the sample is representative of the German population in these respects. To enhance the quality of the sample, individuals, who have answered the survey in an extremely short time ("speeders") have been excluded.

3.2. Measurement

The dependent variable in our research design is formal volunteering. Building on prior surveys like the Swiss Volunteering Survey (Freitag et al., 2016) we measure formal volunteering using the following question:

> We are now interested in all your voluntary activities that you carry out for an association or organization. We are interested in voluntary tasks and work that you carry out unpaid or for a small expense allowance. Have you carried out one or more such activities in the last four weeks?

We create a dichotomous variable based on this question that indicates whether a respondent has at least carried out one voluntary activity within the last four weeks (1 = applicable, 0 = not applicable). Figure 1 shows that around one-third of our respondents actively engaged in volunteering.
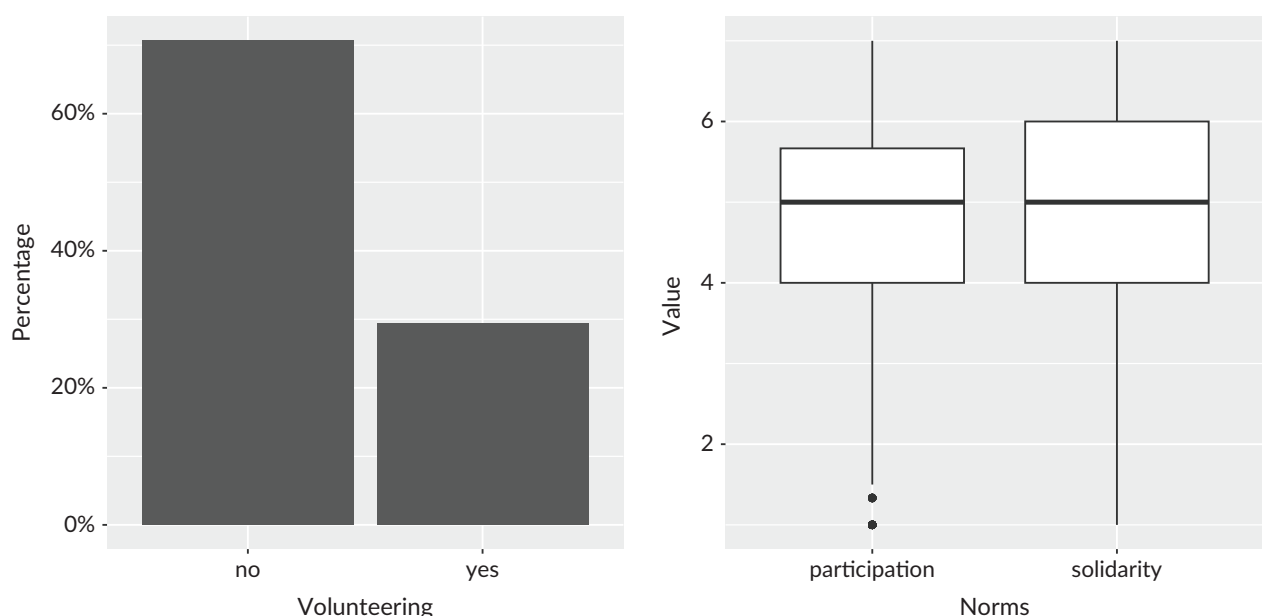
Figure 1. Distribution of the main variables. Source: Data derived from the EXPSOLIDARITY population survey.

Our main independent variables are measures of civic norms. We use the question proposed by Dalton (2008, 2021) to capture the different categories of civic norms:

> There are different opinions about what makes a good citizen. What do you think: To what extent the following things are important to be a good citizen? That someone…
>
> (a) …always votes in elections.
>
> (b) …never tries to evade taxes.
>
> (c) …always obeys laws and regulations.
>
> (d) …keeps watch on the actions of the government.

(e) …is active in social or political associations.

(f) …tries to understand the position of people with other opinions.

(g) …chooses products due to political, ethical or environmental reasons, even if they're more expensive.

(h) …helps people who are badly off in their own country.

(i) …helps people who are badly off in other countries.

Respondents rated their answers on a Likert scale ranging from 1 (*not important at all*) to 7 (*absolutely important*). Since our primary focus is on participation and solidarity norms, we construct two summary indices to capture these categories. The participation index is based on items (a), (e), and (g), while the solidarity index is derived from items (h) and (i). Figure 1 shows the distribution of the two indices. The median value is 5 for both of them; the dispersion is slightly higher for norms of solidarity which show an interquartile range from the scale values 4 to 6.

In addition, our empirical models include moral and socio-structural variables as control and moderating factors. Age is measured as a continuous variable, capturing respondents' age in years. Gender is coded dichotomously (0 = male, 1 = female). Education is categorized into three levels: low (primary education), middle (secondary education), and high (tertiary education). To measure religiosity, we use the centrality of religiosity scale (Huber & Huber, 2012), which assesses religiosity independently of denominational affiliation and captures the importance that religion and religious practice has in someone's life. Specifically, we apply the version of the scale, which consists of seven items covering different dimensions of religiosity, also including meditation and non-theistic elements. It consists of the following questions that are combined into an index of religiosity:

1. How often do you think about religious issues?
2. To what extent do you believe that God or something divine exists?
3. How often do you take part in religious services?
4a. How often do you pray?
4b. How often do you meditate?
5a. How often do you experience situations in which you have the feeling that God or something divine intervenes in your life?
5b. How often do you experience situations in which you have the feeling that you are in one with all?

Answers were captured on a 5-point scale: not at all, not very much, moderately, quite a bit, and very much so (for item (2)) respectively never, rarely, occasionally, often, and very often). To generate the index of religiosity, the mean value of five items is calculated. Therefore, the highest value of (4a) and (4b) is taken as well as the highest value of (5a) and (5b). The index captures religiosity as a metric scale.

3.3. Method

To analyze the relationship between civic norms and volunteering, we estimate logistic regression models. Given that the correlation between our two main independent variables—norms of participation and norms of solidarity—is approximately 0.6 (see Supplementary File, Figure S1), we estimate separate models for each independent variable. To enhance model estimation and facilitate the interpretation of the effects, all continuous variables are mean-centered.

4. Norms of Participation and Solidary Support Volunteering

Turning to the results of our empirical analysis, Figure 2 presents the average marginal effects derived from our two main regression models. In both models, volunteering serves as the dependent variable. Model 1 includes norms of participation as the main independent variable, while Model 2 includes norms of solidarity. At least three key conclusions can be drawn from the regression results. First, norms play a significant role in volunteering. A one-point increase in the participation norms index is associated with a ten-percentage-point increase in the likelihood of volunteering. The effect of solidarity norms is slightly weaker, with a one-point increase in the solidarity norms index corresponding to a six-percentage-point increase in volunteering propensity. Figure 3 presents the predicted probabilities of volunteering over the course of the two main explanatory variables. Other continuous variables in the model are set to their mean and dichotomous variables are fixed at the reference category. The left figure illustrates that an increase over the full range of support for participation norms (minimum to maximum) is related to an increase in the predicted probability of volunteering by almost 25 percentage points. An increase over the full range of support for solidarity norms is related to an increase in the predicted probability of volunteering by about 13 percentage points, as the right figure shows. Again, this indicates that participation norms are substantially more important for volunteering than solidarity norms. Second, although civic norms are important, they are surpassed by religiosity and education as value and socio-structural foundations of volunteering. The effects of these factors exceed those of civic norms in both models, suggesting that religiosity and education are stronger correlates of voluntary engagement. Third, regarding the additional control variables, we find that age is not significantly related to volunteering in our sample, while men exhibit a higher likelihood of volunteering than women. This indicates that formal volunteering in associations and organizations remains a form of civic engagement that is more appealing to men than to women. In conclusion, these initial findings emphasize the overall importance of norms of participation and norms of solidarity for volunteering, while also highlighting the dominant role of religiosity and education in shaping voluntary engagement.

Alternative model specifications support the discussed findings and indicate that norms of participation are more important for the propensity to volunteer than norms of solidarity (see Supplementary File, Table S1). If both norms are included in one model (Model 3), only the link between norms of participation and volunteering remains positively significant. This finding is confirmed if the entire model of civic norms by Dalton (2021) is included in the regression analysis (Model 4). Interestingly, this analysis shows that norms of social order are negatively related to volunteering. The correlation between norms of autonomy and volunteering is not significant.
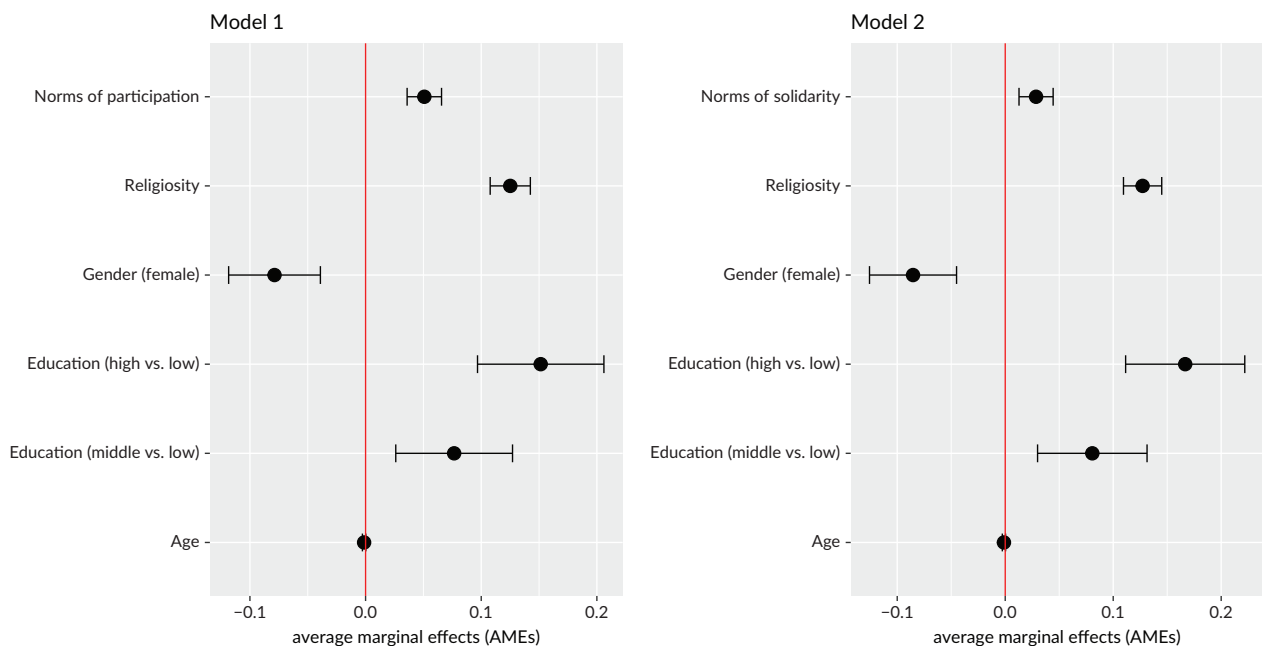
Figure 2. The link between civic norms and volunteering. Notes: The figure shows average marginal effects based on the logistic regression models presented in Table S1; black points indicate the AME and horizontal lines indicate the 95% confidence interval. Source: Data derived from the EXPSOLIDARITY population survey.
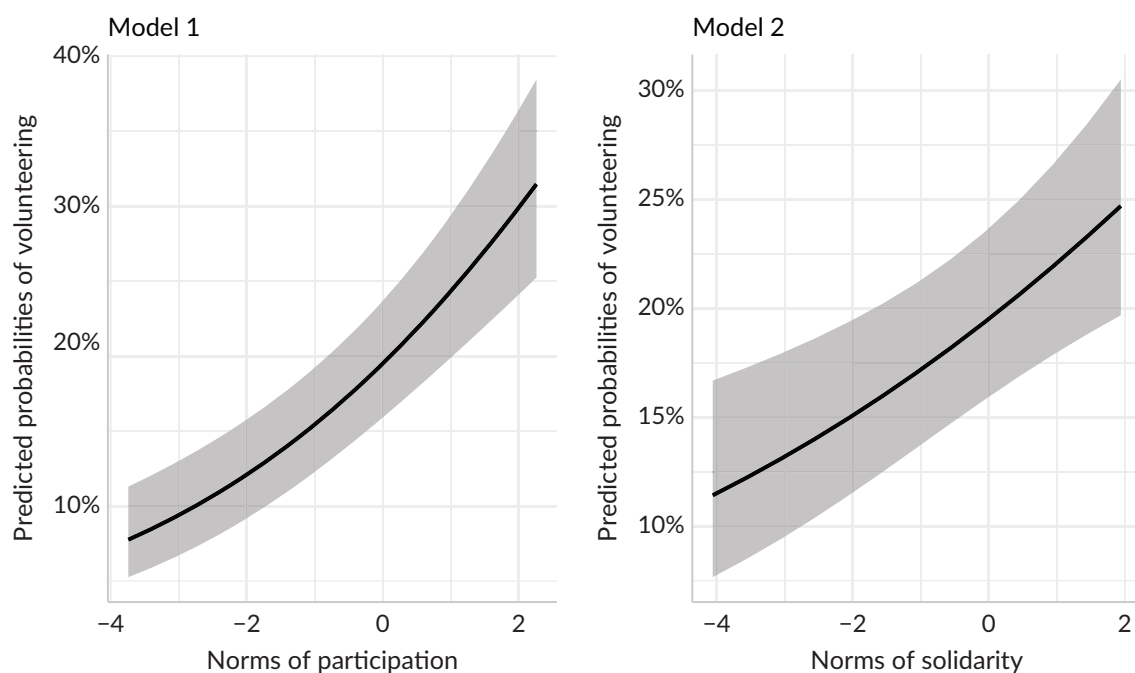
Figure 3. Civic Norms and Predicted Probabilities of Volunteering. Notes: The figure shows predicted probabilities of volunteering across the range of civic norms based on the logistic regression models presented in Table S1. Black lines indicate the prediction, grey areas indicate the 95% confidence interval. Source: Data derived from the EXPSOLIDARITY population survey.

5. Additional Analyses

In a series of additional analyses, we examine whether the relationship between civic norms and volunteering varies across different societal groups. Given the role of socio-demographic characteristics and values for both, norm conformity and volunteering, it is plausible that these factors moderate the link between civic norms and volunteering. This means that the strength and significance of this relationship could vary depending on individual characteristics.

Overall these analyses are rather explorative. Building on existing literature on norm conformity, we have reasons to expect age and gender to be relevant moderators. Regarding age, research suggests that norm conformity tends to decrease with age (Foulkes et al., 2018; Pasupathi, 1999), as older individuals appear to be less susceptible to social influence and peer pressure. Consequently, participation and solidarity norms may have a stronger association with volunteering among younger cohorts. Gender may also influence norm-compliant behavior, particularly in the context of volunteering. Women are more strongly socialized into caring and volunteer work due to both subjective self-perception and structural conditions, reflecting individual values and societal expectations. Therefore, it can be expected that women who strongly endorse participation and solidarity norms are more likely to volunteer than men (Gerstel & Gallagher, 1994; Greeno & Maccoby, 1986). In addition, we test whether education and religiosity moderate the relationship between civic norms and volunteering. Our theoretical expectations for these two factors are less clear-cut and the analyses are rather explorative.

The results are presented in Table 2. Again, separate models are estimated for participation norms (Model 5 to Model 8) and solidarity norms (Model 9 to Model 12). To enhance interpretability, we also estimated separate models for each expected interaction. However, all interaction coefficients are insignificant. This indicates that the relationship between civic norms and volunteering is not influenced by moral or socio-structural factors. While the distribution of norms may vary across groups, their effect on volunteering remains consistent. The positive impact of civic norms on voluntary activities holds across all groups.

Table 2. Civic norms, volunteering, and the moderating role of morality and social structure.

	Model 5	Model 6	Model 7	Model 8	Model 9	Model 10	Model 11	Model 12
(Intercept)	−2.81*** (0.20)	−2.81*** (0.20)	−2.81*** (0.20)	−2.83*** (0.20)	−2.78*** (0.20)	−2.80*** (0.20)	−2.79*** (0.20)	−2.80*** (0.20)
Norms of participation	0.28*** (0.04)	0.28*** (0.06)	0.29*** (0.09)	0.36** (0.12)				
Norms of solidarity					0.15*** (0.04)	0.20** (0.06)	0.15 (0.09)	0.18 (0.11)
* Age	−0.00 (0.00)				−0.00 (0.00)			
* Gender		0.01 (0.09)				−0.10 (0.09)		
* Education (middle)			−0.02 (0.11)				−0.07 (0.12)	
* Education (high)			−0.00 (0.11)				0.08 (0.11)	
* Religiosity				−0.03 (0.04)				−0.01 (0.04)
Age	−0.01 (0.00)	−0.01 (0.00)	−0.01 (0.00)	−0.01 (0.00)	−0.01 (0.00)	−0.01 (0.00)	−0.01 (0.00)	−0.01 (0.00)
Gender	−0.44*** (0.11)	−0.44*** (0.12)	−0.44*** (0.11)	−0.44*** (0.11)	−0.46*** (0.11)	−0.45*** (0.11)	−0.46*** (0.11)	−0.47*** (0.11)
Education (middle)	0.46** (0.16)	0.46** (0.16)	0.46** (0.16)	0.45** (0.16)	0.48** (0.16)	0.48** (0.16)	0.48** (0.16)	0.47** (0.16)
Education (high)	0.85*** (0.16)	0.85*** (0.16)	0.85*** (0.16)	0.84*** (0.16)	0.91*** (0.16)	0.92*** (0.16)	0.91*** (0.16)	0.91*** (0.16)
Religiosity	0.70*** (0.06)	0.70*** (0.06)	0.70*** (0.06)	0.70*** (0.06)	0.69*** (0.06)	0.69*** (0.06)	0.69*** (0.06)	0.70*** (0.06)
AIC	1927.99	1928.08	1930.05	1927.51	1944.95	1945.60	1946.69	1946.76
BIC	1971.87	1971.96	1979.41	1971.39	1988.75	1989.40	1995.97	1990.56
Log Likelihood	−956.00	−956.04	−956.02	−955.76	−964.47	−964.80	−964.34	−965.38
Deviance	1911.99	1912.08	1912.05	1911.51	1928.95	1929.60	1928.69	1930.76
Num. obs.	1781	1781	1781	1781	1764	1764	1764	1764

Notes: Levels of significance: *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$. Source: Data derived from the EXPSOLIDARITY population survey.

6. Conclusion

This study enhances our understanding of voluntary engagement by examining the role of civic norms. While the socio-structural and sociodemographic determinants of volunteering are well-established, the influence of norms in this context remains surprisingly underexplored. Drawing on theoretical frameworks

from psychology, we posit that norms exert strong behavioral effects through processes of socialization, social learning, and self-consistency.

Using data from a representative population survey conducted in Germany, our empirical analyses demonstrate that norms of participation and solidarity are significantly positive related to voluntary engagement. Our findings confirm a stable link between civic norms and volunteering, with participation norms exhibiting a stronger association than solidarity norms. This supports the idea that individuals who internalize participatory and solidarity-based responsibilities as part of their civic identity are more likely to engage in voluntary work. However, our results also highlight that religiosity and education play an even greater role in shaping volunteering behavior than civic norms. When examining group differences, our analyses reveal that the relationship between civic norms and volunteering remains consistent across different societal groups. No significant interaction effects were found between civic norms and moral or socio-structural factors. While the distribution of civic norms may vary, their influence on volunteering is independent of factors such as age, gender, education, or religiosity. This indicates that civic norms function as a universal rather than group-specific foundation for volunteering.

At least three findings should be highlighted and might immediately lead to new research questions. First, it is important to consider the relative substance of the relationship between civic norms and volunteering. Our findings show a consistent positive relationship that one could interpret as norms being an important motivational basis for volunteering. At the same time, education and religiosity seem to be even more important. This hints at the importance of values that are probably even more fundamental than norms, as well as the importance of civic and factual knowledge that individuals acquire through education. Future research should test whether motivation, values, and knowledge are the mechanisms that are at play to better understand the complex relationship between civic norms, religiosity, education, and volunteering. Second, the substantive effect of norms of participation is bigger than the effect of norms of solidarity. Thus, to motivate individuals to volunteer, it might be particularly important to strengthen the participatory nature of it. For many, volunteering seems to be not only an act of solidarity and selfless charity, but also a means of self-realization and the development of self-efficacy. Associations and organizations should therefore give their volunteers the opportunity to shape the structures in which they carry out their voluntary work. Third, one could argue that the insignificance of the interaction effects in our study is a good sign in terms of the effectiveness of civic norms. It means that norms do not have stronger effects for certain groups, like highly educated or highly religious persons. Thus, to foster volunteering as an important aspect of social cohesion, societies should also think about how to strengthen civic norms in the first place. Different societal groups benefit from these norms in similar ways meaning that stronger norms could strengthen civic engagement across all parts of the society.

Although our study yields important findings, its shortcomings should be kept in mind and considered when thinking about future developments in this line of research. The evidence presented is based on survey data that can always be affected by social desirability, meaning that respondents answer survey questions in a way they perceive as expected by their peers. This is particularly relevant for behavioral measures, like our dependent variable volunteering, because we do not know whether respondents report actual or desired behavior. Moreover, the evidence we provide is correlation. Thus, we cannot be entirely sure about the causal direction of the relationships that prove to be significant. Based on theoretical arguments, we have good reasons to assume that causality runs from norms to behavior. Yet, we cannot rule out that civic

engagement also fosters civic norms in the respondents using cross-sectional data. Experimental or longitudinal data would be necessary to investigate the issue of causality. Finally, the geographical scope of our study is limited. We present a single-country study using German survey data. It would be desirable to test with comparative data whether our findings travel to other geographic contexts and whether country-specific factors affect the relationships. It is, for instance, plausible that the link between participation and solidarity norms and volunteering is contingent on the welfare state regime or other political institutions in a country (Ackermann et al., 2023). Finally, our survey was conducted shortly after the Covid-19 pandemic. We cannot rule out that this time, in which many associational activities have been restricted by official regulations, impedes our findings. Formal volunteering was not a typical activity for practicing solidarity norms at this time. It would therefore be interesting to see replications of our study with a greater time lag to the pandemic.

In conclusion, our findings underscore the importance of fostering civic norms as a means to strengthen social cohesion and democratic resilience. In light of ongoing crises, like rising political polarization, the erosion of trust in institutions, and the fragmentation of social networks, voluntary engagement can serve as a stabilizing force in democratic societies. Given that participation and solidarity norms enhance civic engagement, democratic societies should support initiatives that cultivate these norms, for instance in the field of civic education.

Conflict of Interests
The authors declare no conflict of interests.

Data Availability
The data and the R-script to reproduce the analyses was uploaded to the OSF-repository of the corresponding author: https://osf.io/3psfy

LLMs Disclosure
For language editing, we made use of the following LLMs: ChatGPT-4o mini and DeepL.

Supplementary Material

Supplementary material for this article is available online in the format provided by the author (unedited).

References

Ackermann, K. (2019). Predisposed to volunteer? Personality traits and different forms of volunteering. *Nonprofit and Voluntary Sector Quarterly*, *48*(6), 1119–1142. https://doi.org/10.1177/089976401984 8484

Ackermann, K., Erhardt, J., & Freitag, M. (2023). Crafting social integration? Welfare state and volunteering across social groups and policy areas in 23 European countries. *KZfSS Kölner Zeitschrift für Soziologie und Sozialpsychologie*, *75*(Suppl 1), 283–304.

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, *50*(2), 179–211. https://doi.org/10.1016/0749-5978(91)90020-T

Bicchieri, C. (2006). *The grammar of society: The nature and dynamics of social norms.* Cambridge University Press.

Bicchieri, C. (2016). *Norms in the wild: How to diagnose, measure, and change social norms.* Oxford University Press.

Bicchieri, C., & Xiao, E. (2009). Do the right thing: But only if others do so. *Journal of Behavioral Decision Making*, *22*(2), 191–208. https://doi.org/10.1002/bdm.621

Bursztyn, L., & Jensen, R. (2017). Social image and economic behavior in the field: Identifying, understanding, and shaping social pressure. *Annual Review of Economics*, *9*(1), 131–153. https://doi.org/10.1146/annurev-economics-063016-103625

Cialdini, R. B., Kallgren, C. A., & Reno, R. R. (1991). A focus theory of normative conduct: A theoretical refinement and reevaluation of the role of norms in human behavior. *Advances in Experimental Social Psychology*, *24*, 201–234. https://doi.org/10.1016/S0065-2601(08)60330-5

Cialdini, R. B., & Trost, M. R. (1998). Social influence: Social norms, conformity and compliance. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *The handbook of social psychology* (4th ed., pp. 151–192). McGraw-Hill.

Dalton, R. J. (2008). Citizenship norms and the expansion of political participation. *Political Studies*, *56*(1), 76–98. https://doi.org/10.1111/j.1467-9248.2007.00718.x

Dalton, R. J. (2021). *The good citizen: How a younger generation is reshaping American politics* (3rd ed.). CQ Press.

de Tocqueville, A. (2000). *Democracy in America*. The University of Chicago Press. (Original work published in 1835)

Dekker, P., & Halman, L. (Eds.). (2003). *The values of volunteering: Cross-cultural perspectives*. Springer. https://doi.org/10.1007/978-1-4615-0145-9

Denters, B., Gabriel, O., & Torcal, M. (2007). Norms of good citizenship. In J. W. van Deth, J. Ramón Montero, & A. Westholm (Eds.), *Citizenship and involvement in European democracies* (1st ed.). Routledge. https://doi.org/10.4324/9780203965757

Diamond, E. P., & Skrzypek, A. (Eds.). (2024). *The politics of polycrisis*. Dietz.

Durkheim, É. (2019). *Über soziale Arbeitsteilung: Studie über die Organisation höherer Gesellschaften* (8th ed.). Suhrkamp. (Original work published 1893)

Engel, C., & Kurschilgen, M. (2020). The fragility of a nudge: The power of self-set norms to contain a social dilemma. *Journal of Economic Psychology*, *81*, Article 102293. https://doi.org/10.1016/j.joep. 2020.102293

Fonseca, X., Lukosch, S., & Brazier, F. (2019). Social cohesion revisited: A new definition and how to characterize it. *Innovation: The European Journal of Social Science Research*, *32*(2), 231–253. https://doi.org/10.1080/13511610.2018.1497480

Foulkes, L., Leung, J. T., Fuhrmann, D., Knoll, L. J., & Blakemore, S. J. (2018). Age differences in the prosocial influence effect. *Developmental Science*, *21*(6), Article 12666. https://doi.org/10.1111/desc.12666

Freitag, M., Manatschal, A., Ackermann, K., Ackermann, M., Niederberger, L., & Sacchi, S. (2016). *Schweizer Freiwilligen-Monitor 2016* [Data set]. FORS—Swiss Center for Expertise in Social Research. https://doi.org/10.23662/FORS-DS-795-1

Gavrilets, S., & Richerson, P. J. (2017). Collective action and the evolution of social norm internalization. *Proceedings of the National Academy of Sciences*, *114*(23), 6068–6073. https://doi.org/10.1073/pnas.1703857114

Gerstel, N., & Gallagher, S. (1994). Caring for Kith and Kin: Gender, employment, and the privatization of care. *Social Problems*, *41*(4), 519–539. https://doi.org/10.2307/3096987

Giguère, B., Lalonde, R. N., & Taylor, D. M. (2014). Drinking too much and feeling bad about it? How group identification moderates experiences of guilt and shame following norm transgression. *Personality and Social Psychology Bulletin*, *40*(5), 617–632. https://doi.org/10.1177/0146167214521836

Greeno, C. G., & Maccoby, E. E. (1986). How different is the 'different voice'? *Signs*, *11*(2), 310–316.

Grimalda, G., Pondorfer, A., & Tracer, D. P. (2016). Social image concerns promote cooperation more than altruistic punishment. *Nature Communications*, *7*(1), Article 12288. https://doi.org/10.1038/ncomms12288

Gross, J., & Vostroknutov, A. (2022). Why do people follow social norms? *Current Opinion in Psychology*, *44*, 1–6. https://doi.org/10.1016/j.copsyc.2021.08.016

Henriksen, L. S., Koch-Nielsen, I., & Rosdahl, D. (2008). Formal and informal volunteering in a Nordic context: The case of Denmark. *Journal of Civil Society*, *4*(3), 193–209. https://doi.org/10.1080/17448680802559685

Horne, C. (2003). The internal enforcement of norms. *European Sociological Review*, *19*(4), 335–343. https://doi.org/10.1093/esr/19.4.335

Huber, S., & Huber, O. W. (2012). The centrality of religiosity scale (CRS). *Religions*, *3*(3), 710–724. https://doi.org/10.3390/rel3030710

Keizer, K., Lindenberg, S., & Steg, L. (2008). The spreading of disorder. *Science*, *322*(5908), 1681–1685. https://doi.org/10.1126/science.1161405

Knack, S., & Keefer, P. (1997). Does social capital have an economic payoff? A cross-country investigation. *The Quarterly Journal of Economics*, *112*(4), 1251–1288. https://doi.org/10.1162/003355300555475

Lee, Y., & Brudney, J. L. (2012). Participation in formal and informal volunteering: Implications for volunteer recruitment. *Nonprofit Management and Leadership*, *23*(2), 159–180. https://doi.org/10.1002/nml.21060

Lindström, B., Jangard, S., Selbing, I., & Olsson, A. (2018). The role of a "common is moral" heuristic in the stability and change of moral norms. *Journal of Experimental Psychology*: *General*, *147*(2), 228–242. https://doi.org/10.1037/xge0000365

Musick, M. A., & Wilson, J. (2008). *Volunteers: A social profile*. Indiana University Press.

Pasupathi, M. (1999). Age differences in response to conformity pressure for emotional and nonemotional material. *Psychology and Aging*, *14*(1), 170–174. https://doi.org/10.1037//0882-7974.14.1.170

Putnam, R. D. (1995). Bowling alone: America's declining social capital. *Journal of Democracy*, *6*(1), 65–78. https://doi.org/10.1353/jod.1995.0002

Putnam, R. D. (2001). *Bowling alone: The collapse and revival of American community*. Simon & Schuster.

Putnam, R. D., Leonardi, R., & Nonetti, R. Y. (1994). *Making democracy work: Civic traditions in modern Italy*. Princeton University Press.

Stolle, D. (2003). The sources of social capital. In M. Hooghe & D. Stolle (Eds.), *Generating social capital: Civil society and institutions in comparative perspective* (pp. 19–42). Palgrave Macmillan.

Tremewan, J., & Vostroknutov, A. (2021). An informational framework for studying social norms. In A. Chaudhuri (Ed.), *A research agenda for experimental economics* (pp. 19–42). Edward Elgar Publishing. https://doi.org/10.4337/9781789909852.00008

Van der Meer, T. W. G., & Van Ingen, E. J. (2009). Schools of democracy? Disentangling the relationship between civic participation and political action in 17 European countries. *European Journal of Political Research*, *48*(2), 281–308. https://doi.org/10.1111/j.1475-6765.2008.00836.x

van Deth, J. W. (2009). Norms of citizenship. In R. J. Dalton & H. Klingemann (Eds.), *The Oxford handbook of political behavior* (1st ed., pp. 402–417). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199270125.003.0021

Walzer, M. (1996). *Zivile Gesellschaft und amerikanische Demokratie*. Fischer Taschenbuch Verlag.

Wang, L., Mook, L., & Handy, F. (2017). An empirical examination of formal and informal volunteering in Canada. *VOLUNTAS: International Journal of Voluntary and Nonprofit Organizations*, *28*(1), 139–161. https://doi.org/10.1007/s11266-016-9725-0

Wilkinson, R. G., & Pickett, K. (2010). *The spirit level: Why equality is better for everyone.* Penguin Books.

Wilson, J. (2000). Volunteering. *Annual Review of Sociology*, *26*(1), 215–240. https://doi.org/10.1146/annurev.soc.26.1.215

Wilson, J. (2012). Volunteerism research: A review essay. *Nonprofit and Voluntary Sector Quarterly*, *41*(2), 176–212. https://doi.org/10.1177/0899764011434558

Zmerli, S. (2010). Social capital and norms of citizenship: An ambiguous relationship? *American Behavioral Scientist*, *53*(5), 657–676. https://doi.org/10.1177/0002764209350829

About the Authors

Kathrin Ackermann is a professor of political science at the University of Siegen. Her fields of research include empirical democracy research, political participation, public opinion, and social capital. She has widely published on volunteering, attitudes towards democracy and the psychological foundations of public opinion and political behavior.

Jonathan Mylius studied journalism and political science and is now studying theology in Berlin. He has been involved as research assistant in the EXPSOLIDARITY project at Heidelberg University. His areas of interest are evangelicalism and the sociology of religion, currently focusing on the influence of evangelicalism on democratic culture.

Annette Haussmann is a professor of practical theology with a focus on pastoral and spiritual care at Heidelberg University. She is also director of the Center for Pastoral Care at the regional Church of Baden (Germany) and a trained psychotherapist for CBT. Her current research includes mental health and religion, digital spiritual care, and psychology of religion.

Stefanie Wiloth is a gerontologist. She completed her PhD at Heidelberg University and conducted research at the Institute for Gerontology and the Institute for Diaconal Sciences. Since August 2024, she has been Head of the Competence Centre for Social Innovation and Ageing at the Institute for Ageing Research at the Eastern Switzerland University of Applied Sciences in Switzerland.

cogitatio

Breaking False Polarization: How Information on Descriptive Norms Mitigates Worry Rooted in Polarization (Mis)perceptions

Tom Nijs 🔘

ERCOMER, Utrecht University, The Netherlands

Correspondence: Tom Nijs (t.nijs@uu.nl)

Abstract

Worry about polarization in society, particularly around the topic of immigration, is widespread despite the lack of substantial evidence supporting the existence of actual polarization of attitudes. This study explores whether this widespread "polarization panic" can be attributed to misperceptions of the descriptive norm, more specifically, to overestimations of polarization in society, a phenomenon known as false polarization. I investigated whether Dutch participants were more worried about polarization when they perceived stronger polarization in immigration attitudes due to a misperception of attitudinal extremity as the descriptive norm and whether correcting their misperceptions with accurate information about the actual descriptive norm reduced this association. A pre-registered survey-embedded experiment ($N = 925$) revealed that the significant positive relationship between perceptions of polarization and polarization worry disappeared when participants were provided with accurate information about the descriptive norm in society. However, this effect was only observed among participants who realized and acknowledged that they overestimated the differences in attitudes. These results suggest that during times of widespread media reports on alarming increases in polarization, informing individuals about the actual descriptive norm can alleviate worry amongst those who overestimate polarization. This approach could potentially facilitate respectful dialogue about the hotly debated topic of immigration. Nevertheless, the effectiveness of this strategy hinges on ensuring that the descriptive norm is correctly interpreted, leading individuals to realize that their worry was based on misperceptions.

Keywords

descriptive norms; false polarization; perceived polarization; polarization worry

---

1. Introduction

Polarization is widely regarded as one of the most severe societal issues globally. The World Economic Forum (2024) ranked societal polarization as the third most critical short-term global risk in 2024 and three in ten Americans consider it a top issue facing their country (Skelley & Fuong, 2022). Researchers argued that this widespread "polarization panic" is surprising (Miltenburg et al., 2022; Muis, 2024), as evidence for significant attitudinal polarization is lacking in many societies (Baldassarri & Bearman, 2007; Dekker, 2022). It has been suggested that "polarization is perceived rather than actually occurring" (Muis, 2024, p. 273) and that "there is more agreement on many topics than people think" (NOS, 2022), implying that polarization worry stems from misperceptions of the prevalence of attitudinal extremity in society (i.e., the descriptive norm), leading individuals to overestimate polarization. Although this phenomenon, known as false polarization, is robustly found in the US context (Levendusky & Malhotra, 2016), I am not aware of any studies empirically testing this in the European context, nor of any that have tested the relationship between perceptions of polarization and polarization worry.

Using a pre-registered survey-embedded experiment ($N = 925$), I tested the implicit claim that polarization worry is rooted in (mis)perceptions of polarization and additionally, whether this relationship can be mitigated by providing information about the actual descriptive norm in society. If individuals are worried about polarization because they have an inaccurate perception of polarization, offering accurate information on the real attitudes in society may change their polarization perception so that their initial perception will become less predictive of polarization worry.

The hypotheses were tested in relation to immigration attitudes in the Netherlands. Polarization spans many topics, but immigration is consistently identified as one of the most polarized issues in Europe. Immigration was indicated as the issue most likely to create a division in society by European citizens (Herold et al., 2023) and was found to be the most polarized issue amongst European political parties (Reiljan, 2019). I focus on the Netherlands, where polarization worry is widespread, despite little evidence of increasing attitudinal polarization. A majority of Dutch citizens indicated to be very worried about increasing polarization (SIRE, 2023). However, Dekker (2022) concluded in his book on Dutch polarization that "no evidence was provided for (increasing) polarization as a dominant trend" (p. 39). Supporting this, Dutch citizens were found to be among the least polarized in Europe (Herold et al., 2023). In this context, where evidence for increasing attitudinal polarization is lacking, understanding the roots of polarization worry, and exploring ways to reduce it is particularly relevant.

2. Polarization Worry

Polarization has been defined in various ways (Fernbach & Van Boven, 2022) and different types of polarization have been studied in the academic literature. Although affective polarization—referring to increasing dislike and distrust between those with differing views—has recently received increasing attention (see Iyengar et al., 2019; Levendusky, 2018; Orhan, 2022), researchers have traditionally studied attitudinal or ideological polarization, which refers to increasing divergence among people in their political beliefs, policy preferences, and ideological positions (see Abramowitz & Saunders, 2008; Fiorina & Abrams, 2008). Polarization has raised widespread worry across the globe, a phenomenon dubbed "polarization panic" (Muis, 2024; Ros, 2023), which can have several harmful implications. Similar to the negative mental

health impacts linked to worry about other societal issues like climate change (Clayton, 2020) and the Covid-19 pandemic (Hossain et al., 2020), polarization worry is suggested to have detrimental consequences for mental health (Nayak et al., 2021; Smith, 2022; Smith et al., 2019).

Moreover, polarization worry might make individuals withhold their views, to avoid confrontations. According to the Spiral of Silence theory (Noelle-Neumann, 1974), individuals withhold opinions that they believe others disagree with, to prevent being socially isolated (Matthes et al., 2017; Wuestenenk et al., 2025). In line with this, Americans were found to self-censor their own views more in times of stronger affective polarization (Gibson & Sutherland, 2023). Polarization worry might therefore impede dialogue about contentious subjects, potentially harming the functioning of democracies (Carpini et al., 2004; Dahlgren, 2002). This means that, in line with previous arguments that actual polarization can harm democratic functioning (Graham & Svolik, 2020; Webster & Abramowitz, 2017), merely worrying about polarization—regardless of whether that worry is justified—may itself be detrimental to democracy. Another harmful implication is that when individuals worried about polarization retreat into attitudinal bubbles (see Webster & Abramowitz, 2017), they discuss contentious topics with people with similar attitudes only. Literature on group polarization shows that discussions with like-minded individuals lead to more extreme views (Cooper et al., 2001; Myers & Lamm, 1976). Radicalization dynamics can lead individuals with similar opinions to mutually reinforce each other, leading to more radical views over time (Baumann et al., 2020). Therefore, polarization worry can become a self-fulfilling prophecy (Muis, 2024).

Despite these harmful implications, I am unaware of studies systematically explaining polarization worry. Based on the limited and mixed empirical evidence of the existence of attitudinal polarization (e.g., Abramowitz & Saunders, 2008; Baldassarri & Bearman, 2007; Cowan & Baldassarri, 2018; Fischer & Hout, 2006; Lelkes, 2016), researchers have implied that this worry is rooted in misperceptions of attitudinal extremity as the descriptive norm in society, leading individuals to overestimate polarization (Baldassarri & Bearman, 2007; Dekker, 2022; Miltenburg et al., 2022; Muis, 2024). This claim has not been systematically tested.

3. Perceptions of Polarization

Extensive research has revealed that individuals tend to have a biased perception of which behaviours, attitudes, or beliefs are common, referred to as the descriptive norm (Cialdini et al., 1990). For example, individuals consistently misperceive the prevalence of behaviors or attitudes among their peers, a pattern frequently examined in the context of adolescent substance use (Amialchuk et al., 2019; Lintonen & Konu, 2004; Perkins et al., 2011), as well as within the general public at large (Andre et al., 2024; Broockman & Skovron, 2018; Wenzel, 2005). Similarly, Americans were found to systematically overestimate attitudinal extremity in society leading to overestimating attitudinal polarization between Republicans and Democrats (Ahler, 2014; Levendusky & Malhotra, 2016; Westfall et al., 2015). This false polarization bias reflects a misperception of a *descriptive norm*, as it focuses on the perceived prevalence of extreme attitudes among others, distinct from an *injunctive norm* (Cialdini et al., 1990), which would concern the perceived social approval or disapproval of attitudinal extremity.

Often without explicitly testing whether polarization was *overestimated* or *false*, researchers have explored the consequences of perceptions of polarization across various contexts and political issues. Studies have

identified both positive and negative consequences. On the positive side, perceptions of stronger polarization relate to more collective action, voting, and political participation (Enders & Armaly, 2018; Moral, 2017; Roblain & Green, 2021; Westfall et al., 2015). For example, Roblain and Green (2021) found that those who perceive stronger polarization of immigration attitudes regard their immigration attitudes as more central to their self-definition and feel more capable to make a change, prompting collective action. On the negative side, perceptions of stronger polarization are linked to lower political and social trust, more negative affective evaluations of outparties and their candidates, and lower political efficacy (Enders & Armaly, 2018; Lee, 2022). It is argued that perceived polarization decreases a sense of shared values and belonging, reducing trust among members of society (Lee, 2022).

Although studies have shown that a distorted perception of reality (on conspiracy beliefs see, for example, Hornsey et al., 2021) and perceptions of the social norm (on alcohol use see, for example, Labrie et al., 2014) can relate to worry about a range of topics, no studies have examined whether perceptions of attitudinal polarization in society relate to worry about rising polarization. As people fear that strongly opposing beliefs in society will increase conflict, limit free expression, and make countries ungovernable (SIRE, 2023), I test the claim that:

> People who perceive more polarization in immigration attitudes in society are more worried about rising polarization in society (H1).

4. The Mitigating Role of Offering Accurate Information on the Descriptive Norm

When worry about polarization originates from perceptions of polarization that are often incorrect, providing accurate information about the descriptive norm in society may mitigate this relationship. Descriptive social norms interventions traditionally aimed to promote specific attitudes and behaviours by informing individuals about others who hold similar attitudes or engage in similar behaviours (Gifford & Nilsson, 2014; Larimer & Neighbors, 2003; Niemiec et al., 2020; Rand & Yoeli, 2024). Consistent with this, Peters (2021) argues that repeated reports by researchers and journalists about rising polarization may exacerbate polarization by conveying a descriptive norm of attitudinal extremity. Other scholars have argued that since individuals tend to overestimate attitudinal extremity, providing accurate information about actual levels of polarization can lead to the adoption of less extreme attitudes (Ahler, 2014; Blatz, 2024; Lees & Cikara, 2021).

In the current study, I do not test the traditional norm conformity hypothesis that offering accurate information on actual levels of polarization affects attitudinal extremity. Instead, I propose that individuals who receive accurate information will change their polarization perception so that their initial perception will become less predictive of polarization worry. This applies both to individuals who overestimate and underestimate polarization. Although it has been robustly found that many tend to overestimate polarization (Ahler, 2014; Levendusky & Malhotra, 2016; Westfall et al., 2015), others may underestimate it or have accurate perceptions. Providing truthful information about the descriptive norm in society functions as a self-reflection exercise, helping individuals recognize that their (lack of) worry was based on misperceptions of reality (Blatz, 2024). For those who overestimate polarization, accurate information reveals that attitudinal extremity is not the descriptive norm in society, leading to reduced polarization worry. Consistent with this, prior research showed that offering accurate information on levels of polarization reduced perceived attitudinal extremity (Blatz, 2024; Fernbach & Van Boven, 2022). For those who underestimate

polarization, for example because they are disengaged from politics and unaware of attitudinal divides in society, such information may lead to the realization that attitudinal differences are real and widespread, potentially increasing polarization worry. Thus, the effects of providing accurate information and participants' initial perceptions are strongly interdependent, resulting in the following hypothesis:

> The positive relationship between perceived polarization and polarization worry is weaker when people are informed about the actual levels of polarization (H2).

5. Methods

5.1. Sample

I surveyed 1,010 Dutch participants who indicated that they and both their parents were born in the Netherlands via the online platform Panel Inzicht in February and March 2024. Although the exclusion of participants with a migration background was not necessary for testing the hypotheses, the questionnaire included several other items and experiments to test hypotheses unrelated to the current research. Some of these hypotheses could not be tested on participants with a migration background, which is why they were not included in the data collection. The study was approved by the Ethics Review Board of the Faculty of Social & Behavioural Sciences at Utrecht University. Before data collection, I calculated that a regression analysis with seven predictors, assuming a power of .8 and a small effect size, required a sample of 725 participants. I considered a small effect size since effect sizes in moderation tests are generally smaller (Aguinis et al., 2005). Seventy-three participants were excluded because they failed an attention check at the start of the questionnaire. An additional twelve participants were excluded because they had missing values on at least one of the items used in the analyses. The final sample ($N = 925$) was diverse in terms of gender (53.5% women), age (18–89, $M = 48.6$, $SD = 18.1$), and education level (14% low secondary school or less, 39% high school or vocational training, and 47% [applied] university).

5.2. Procedure

After giving informed consent and indicating background characteristics, participants were offered a scale used in the Dutch Parliamentary Election Study (DPES) 2023 to measure immigration attitudes. The scale taps into assimilation beliefs and asks whether migrants should be able to live in the Netherlands while retaining their own culture or must adapt completely to Dutch culture (see Section 5.3). Participants used this scale to indicate their own attitudes and the attitudes of the electorates of four Dutch political parties to measure perceptions of polarization. Subsequently, participants were randomly divided in either a treatment ($n = 460$) or control ($n = 465$) condition. In the treatment condition, participants were shown the following (translated from the Dutch): "The Dutch Parliamentary Election Study 2023 shows that the average voters for those parties answered the question as follows…" Subsequently, a figure was shown in which the results of the DPES 2023 were visualized (see Figure 1). To make sure that participants actively compared their own perception to the actual attitudes of the four electorates as presented in the figure, participants in the treatment condition were asked: "When you look at this picture, are the differences in opinions bigger or smaller than you expected?" The aim of this question was to encourage participants to actively reflect upon whether or not this new information aligned with their initial perception of polarization. People in the control condition did not receive any information and were not asked this question. Directly after
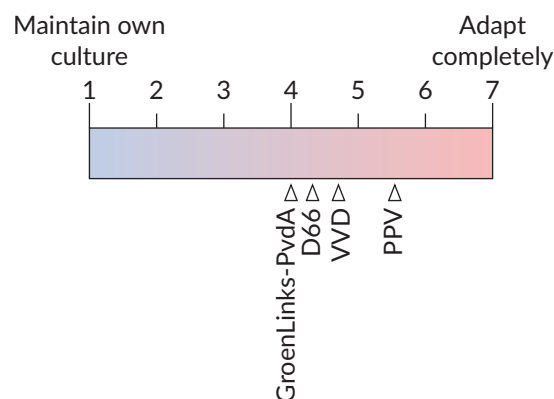
Figure 1. Figure shown in the treatment condition of the experimental manipulation (translated from the Dutch), based on the actual attitudes of electorates found in the DPES 2023.

the experimental manipulation, participants answered an item measuring polarization worry, used as the dependent variable.

The design was inspired by a manipulation of Ahler (2014), who also visualized descriptive results from nationwide public opinion research and asked participants to reflect upon the information offered. The approach to first ask participants to indicate their perception of attitudes of the electorates and present the actual attitudes of these electorates right after, is inspired by the interactive intervention of Blatz (2024) in which participants were informed that they engaged in false polarization right after reporting their perceptions of attitudes.

5.3. Measures

5.3.1. Polarization Worry

The dependent variable polarization worry was measured with one item treated as a continuous scale: "To what extent do you agree with the following statement? I worry that disagreements about migration are increasing in the Netherlands" (1 = *completely disagree*, 5 = *completely agree*).

5.3.2. Perceived Polarization

To measure the independent variable perceived polarization, participants answered the following questions:

> The question about preserving culture was also presented to a representative sample of Dutch people in 2023 (in the DPES). In your estimation, where do you think the following groups have placed themselves on a scale from 1 to 7, where 1 means preservation of their own culture for people with a migration background and 7 means that they have to adapt completely?

The groups presented were *the average Dutch person*, *the average VVD voter*, *the average GroenLinks-PvdA voter*, *the average PVV voter*, and *the average D66 voter*.

Perceived polarization in multiparty democracies has been measured in various ways, such as assessing the distance between the most ideologically distant political parties (Yang et al., 2016) or between all parties to the left and the right of the mean stance (Moral, 2017). In this study, I included PVV and GroenLinks-PvdA in the measure of perceived polarization as they are the biggest parties in parliament representing strongly conservative and strongly progressive stances on immigration, respectively. I additionally included VVD and D66, which are more moderately conservative and progressive, respectively. These four parties were among the five biggest in the Dutch parliamentary elections in 2023 and played a salient role in election debates on immigration, which voters considered the most important topic in those elections (Voogd et al., 2024). Other big parties, such as NSC and BBB were not considered as they were relatively new, and their immigration stances were expected to be insufficiently salient. Parties with immigration stances that may be considered even more progressive than the stance of GroenLinks-PvdA, such as DENK and BIJ1, were not considered as they were very small and might not be well-known to the general public. Following approaches by, for example, Yang et al. (2016) and Levendusky and Malhotra (2016), I calculated the absolute difference between the average scores of the two more conservative electorates (VVD and PVV) and the average scores of the two more progressive electorates (GroenLinks-PvdA and D66), resulting in a scale ranging between 0 and 6.

5.3.3. Perceived Accuracy of Polarization Perception

Participants assigned to the treatment condition received the following question right after the figure was shown: "When you look at this picture, are the differences in opinions bigger or smaller than you expected?" Answer categories were *bigger than expected*, *approximately as I expected*, and *smaller than expected*. Although this question was primarily included to make participants actively reflect on the information presented in the figure, it was used as a manipulation check in exploratory analyses (see the section Analyses).

5.3.4. Control Variables

I controlled for gender, age, education level, and participants' own immigration attitude. Gender was a dichotomous variable (0 = *male*; 1 = *female*). Age in years and education level (1 = *no education* to 7 = *University degree*) were treated as continuous variables. Immigration attitude was measured with the question: "Where would you place yourself on a scale from 1 to 7, where 1 means preservation of their own culture for people with a migration background and 7 means that they have to adapt completely?"

5.4. Analyses

Before testing the hypotheses, I descriptively examined whether false polarization of immigration attitudes existed among Dutch citizens. To do so, I performed a one-sample *t*-test to compare the mean perceived polarization (measured as the difference between the average perceived attitudes of the PVV and VVD electorates and the average perceived attitudes of the GroenLinks-PvdA and D66 electorates) to the actual polarization as found in the DPES 2023 (measured as the difference between the average self-reported attitudes of the PVV and VVD electorates and the average self-reported attitudes of the GroenLinks-PvdA and D66 electorates).

To test the hypotheses, I performed an Ordinary Least Squares regression moderation model with polarization worry as the dependent variable, perceived polarization as the independent variable, and the

experimental manipulation (1 = *treatment*, 0 = *control*) as the moderator. Note that H1 was tested with the moderation included, as the perceived polarization coefficient reflects the relationship between perceived polarization and polarization worry among people who did not receive the manipulation (the moderator has value 0). All control variables were added in a second model. Unstandardized beta coefficients were reported in all analyses. Additionally, I explored whether the strength of the moderating effect of the experimental manipulation to test H2 was influenced by answers to the question measuring perceived accuracy of polarization perception, that was posed right after the figure in the treatment condition. I tested whether filtering out the participants who indicated that differences in opinions were *approximately as expected* would increase the effect of the moderation. This way, I used the question as a manipulation check, based on the idea that the relationship between perceived polarization and polarization worry is likely not influenced when participants consider their initial perceptions of polarization to be accurate (Ahler, 2014).

All analyses, including the exploratory analyses, were pre-registered prior to data collection (https://osf.io/x28yf). Note that this pre-registration also included hypotheses that were tested in a different article (in the pre-registration, referred to as H1a and H1b). In that article, I examined to what extent attitudinal extremity in the social network explains perceptions of polarization. Perceived polarization was used as a dependent variable there, whereas it was used as an independent variable in the current study. H1 and H2 in the current article were respectively called H2a and H2b in the pre-registration. Note that the analyses in the manuscript slightly deviate from the analysis plan in the pre-registration to keep the results section concise. More specifically, I did not test a model without a moderation, and without own immigration attitude.

6. Results

6.1. Descriptive Statistics

On average, participants in the control condition, who did not receive any information, scored 3.82 (on a scale from 1 to 5) on polarization worry, indicating that they were closest to "agreeing" with the statements that they worry that disagreements about migration are increasing in the Netherlands. Of those participants, 72% *agreed* (4) or *completely agreed* (5) with the statement, showing that polarization worry was widespread in this sample. Participants in the treatment condition, who did receive information about levels of polarization were slightly less worried (M = 3.69) than people in the control condition (t(919) = 2.00, p = .046).

Figure 2 shows how participants on average perceived the attitudes of the four electorates and how the four electorates self-reported their attitudes in the DPES 2023. Whereas the attitude of the VVD electorate was perceived very accurately (as indicated by the proximity of the orange dots), the attitudes of the other electorates were perceived as more extreme than self-reported by the electorates in the DPES 2023. Participants on average perceived the immigration attitudes of the electorates of the two conservative parties (VVD and PVV) on the one hand and the two progressive parties (GroenLinks-PvdA and D66) on the other hand to differ by 2.24 points on a 7-point scale. In the DPES 2023, the self-reported immigration attitudes of the electorates of the two conservative parties on the one hand, and the two progressive parties on the other hand differed only 1.08 points. A one-sample *t*-test suggests that the mean perception of polarization was significantly higher than the actual polarization (t(924) = 24.47, p < .001) suggesting that false polarization of immigration attitudes existed among this Dutch sample.
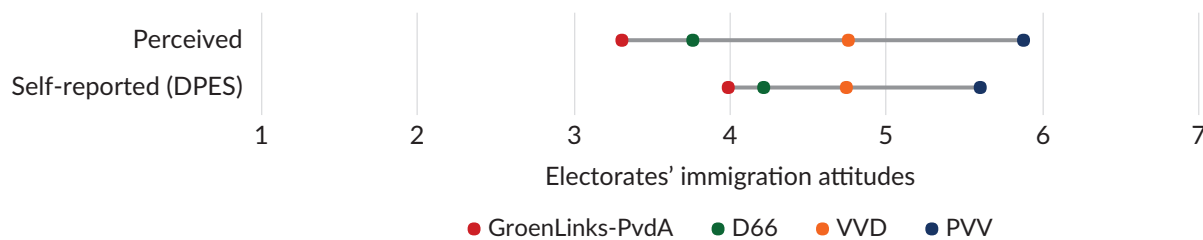
The difference in electorates' immigration attitudes either as perceived by participants in the current dataset, or as self-reported by participants in the DPES 2023.

Figure 3 displays the distribution of perceived polarization. The dashed line represents the self-reported level of polarization based on the DPES 2023. It shows that the vast majority of participants overestimated polarization as 72% scored higher than 1.08 on perceived polarization. Still, a substantial number of participants underestimated or accurately perceived polarization.

Of the participants who received accurate information about polarization in the treatment conditions, 12% (53) indicated that the differences were bigger than expected, 45% (206) indicated that they were approximately as expected, 43% (200) indicated that they were smaller than expected, and one did not respond to this question. This suggests that less than half of the participants in the treatment condition realized that they overestimated polarization. Those who indicated that the differences were smaller than expected did perceive the strongest polarization ($M = 2.97$), compared to those who indicated that they were approximately as ($M = 1.69$) and
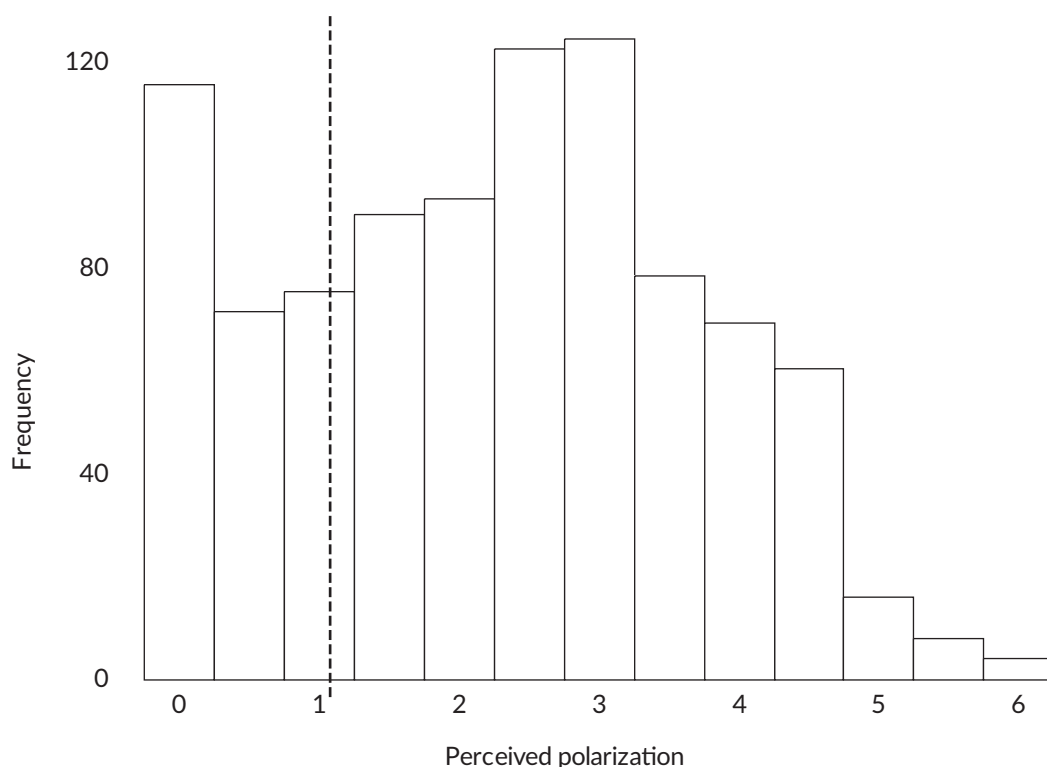
Figure 3. Distribution of perceived polarization. Notes: Perceived polarization was measured as the absolute difference between the average perceived attitudes of the two more conservative electorates and the two more progressive electorates; the dashed line represents the self-reported level of polarization based on the DPES 2023 (1.08).

bigger than expected (M = 1.93). The average own immigration attitude, used as a control variable, was 4.74 (on a scale from 1 to 7), which was significantly more conservative than the neutral midpoint (t(924) = 14.36, $p < .001$).

6.2. Testing the Hypotheses

A moderation model excluding control variables showed that mean centred perceived polarization was significantly and positively related to polarization worry (B = .08, SE = .03, p = .008). Since the moderation of the experimental manipulation was included, this result indicates a positive relationship between perceived polarization and polarization worry among those who did not receive the manipulation, which is in line with H1. See all results in Table 1. The experimental manipulation in which participants were informed about the actual levels of polarization had a significant and negative effect on polarization worry (B = −.13, SE = .06, p = .040), which suggests that the manipulation had a main reducing effect on polarization worry, in line with the t-test reported in the descriptive statistics. However, the manipulation did not significantly moderate the positive association between perceived polarization and polarization worry (B = −.06, SE = .04, p = .181), which is not in line with H2. An R^2 of .01 indicated that only 1% of the variance of polarization worry was explained by the independent variables. The results did not substantially change when control variables were added. Age was the only control variable that was significantly and positively related to polarization worry (B = .01, SE = .00, $p < .001$) indicating that older participants were more worried.

Table 1. Ordinary Least Squares regression models with polarization worry as the dependent variable.

	Main analysis (complete sample)		Exploratory analysis (restricted sample)	
	Excluding control variables	Including control variables	Excluding control variables	Including control variables
	B (SE)	B (SE)	B (SE)	B (SE)
Perceived polarization	.08 (.03)**	.08 (.03)**	.08 (.03)**	.08 (.03)**
Manipulation (1 = accurate information offered)	−.13 (.06)*	−.14 (.06)*	−.08 (.08)	−.09 (.08)
Perceived polarization * manipulation	−.06 (.04)	−.07 (.04)	−.12 (.05)*	−.13 (.05)*
Gender (female)		.06 (.06)		.07 (.07)
Age		.01 (.00)***		.01 (.00)**
Education level		−.01 (.03)		−.00 (.03)
Own immigration attitude		−.02 (.02)		−.02 (.02)
N	925	925	718	718
R^2	.01	.03	.01	.02

Notes: *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$; B represents non-standardized beta coefficients and SE represents standard errors; in the restricted sample, participants were excluded who indicated that differences in opinions were *approximately as expected*.

6.3. Exploring the Role of Perceived Accuracy of Polarization Perception

As pre-registered, I explored whether the strength of the moderating effect of the experimental manipulation to test H2 was stronger when filtering out the participants who indicated that differences in opinions were "approximately as expected" after being informed about the actual levels of polarization. One additional participant who did not respond to this question was also excluded. On this restricted sample ($N = 718$), the experimental manipulation did have a negative and significant moderating effect, both when control variables were excluded ($B = -.12$, $SE = .05$, $p = .027$) and included ($B = -.13$, $SE = .05$, $p = .018$; see also Table 1). The main effect of the experimental manipulation did not reach statistical significance, both when control variables were excluded ($B = -.08$, $SE = .08$, $p = .295$) and included ($B = -.09$, $SE = .08$, $p = .245$).

Simple slopes in Figure 4 suggested that polarization perceptions were positively and significantly related to polarization worry when participants were not informed about the accurate levels of polarization ($B = .08$, $SE = .03$, $p = .010$) but unrelated to polarization worry when individuals were informed ($B = -.05$, $SE = .04$, $p = .296$). This suggests that the positive association between perceptions of polarization and polarization worry can be mitigated by informing participants about actual levels of polarization, which is in line with H2, but only when participants realize and acknowledge that these actual levels of polarization diverge from their initial perceptions of polarization.

To explore these results further, I tested whether the moderating effect of the manipulation existed among those who indicated that the differences were smaller than expected, among those who indicated that the differences were bigger than expected, or among both groups. These exploratory analyses were not
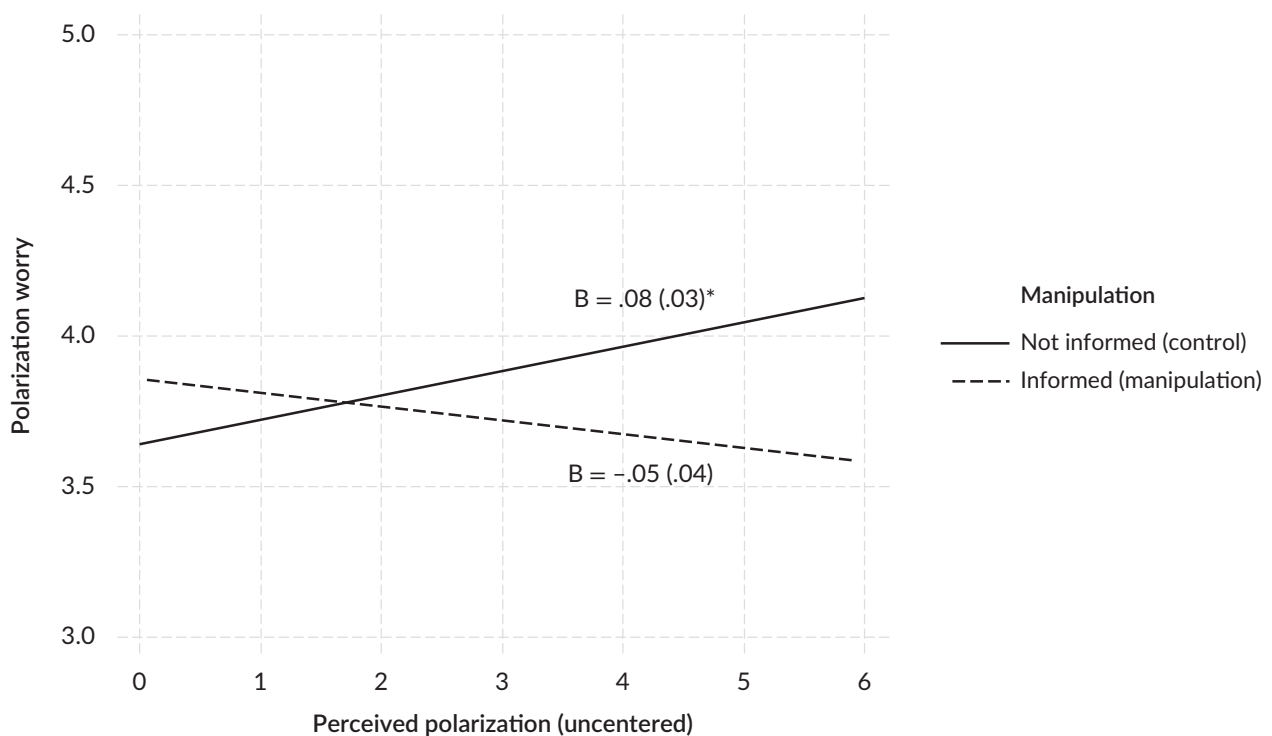

Figure 4. Simple slopes of perceived polarization when individuals were *informed* (1) or *not informed* (0) about actual levels of attitudinal polarization.

pre-registered. Using the full sample, I included three dichotomous moderators, one for those who indicated that the differences were bigger, one for those who indicated they were as expected, and one for those who indicated they were smaller, with participants in the control condition as the reference group. These results, reported in Table S1 in the Supplementary File, suggested that the manipulation only moderated the relationship between polarization perceptions and polarization worry amongst those who indicated that the differences were smaller than expected ($B = -.17$, $SE = .06$, $p = .006$). The manipulation did not moderate this relationship amongst those who indicated that the differences were bigger than expected ($B = -.02$, $SE = .08$, $p = .779$), nor amongst those who indicated that the differences were as expected ($B = -.02$, $SE = .06$, $p = .722$).

7. Discussion and Conclusions

"The only thing we do not seem to be polarizing on these days is the worry that we are polarizing" (Muis, 2024, p. 17). This study provides initial evidence that widespread worry about polarization can, at least in part, be explained by misperceptions of the descriptive norm, more specifically, by overestimations of polarization in society. Dutch participants perceived the differences in immigration attitudes between the electorates of two progressive parties and two conservative parties to be more than twice as large as the differences reported by these electorates themselves. This offers initial evidence of systematic false polarization in the Netherlands, a context where polarization worry is widespread while evidence for attitudinal polarization is lacking (Dekker, 2022; Miltenburg et al., 2022; SIRE, 2023).

Importantly, the findings revealed that the significant positive relationship between perceptions of polarization and polarization worry disappeared when participants were provided with accurate information about the descriptive norm in society. However, this effect was only observed among participants who realized and acknowledged that they overestimated the differences in attitudes. These results suggest that during times of widespread media reports on alarming increases in polarization (Baldassarri & Bearman, 2007; Peters, 2021), informing individuals about the actual descriptive norm may have the potential to alleviate worry amongst those who perceive high levels of polarization. This approach could potentially facilitate respectful dialogue about the hotly debated topic of immigration, particularly between individuals with seemingly divergent views (Noelle-Neumann, 1974). Nevertheless, the effectiveness of this strategy partly depends on individuals correctly interpreting the descriptive norm, leading individuals to realize that their worry was based on misperceptions. Our results suggest that this is not an easy task. Of those who indicated that the actual differences in attitudes between the electorates were approximately as expected, a vast majority still strongly overestimated polarization, suggesting that many participants did not interpret the information as expected. This suggests that numerical results from public opinion research should be accompanied by sufficient explanation and interpretation (for an inspiring example see the intervention from Blatz, 2024).

Although the positive relationship between perceived polarization and polarization worry, and the effect of the manipulation reached statistical significance, the findings should be treated with care. The model accounted for only a very small portion of the variance of the dependent variable. Therefore, I do not claim that perceptions of polarization are *at the root* of polarization worry, nor that offering information about the descriptive norm is a guaranteed solution to alleviate worry among those who perceive high levels of polarization. Polarization perceptions are rather a small piece of the polarization panic puzzle.

One can speculate on other puzzle pieces based on existing literature. Polarization worry might not only be explained by perceptions of polarization amongst the mass public, but also by perceptions of institutionalized polarization amongst political elites, partisan media, and social media (Enders, 2021; Wilson et al., 2020). Media and politicians tend to portray an exaggerated picture of polarization in society which might particularly fuel into polarization worry (see Yang et al., 2016). Literature on elite cueing and media framing has consistently shown that political elites and media can have a strong impact on attitudes and beliefs among the public (see Banks et al., 2020; Robison & Mullinix, 2015). In that case, offering information on the descriptive attitudinal norm amongst the mass public might be less impactful. Rather, Iyengar and Westwood (2014) suggested that it is important that political leaders set an injunctive norm that the expression of hostility towards people with opposing beliefs is unacceptable. This might also alleviate polarization worry. Relatedly, polarization worry might not only be explained by perceptions of attitudinal differences, but also by how people with different attitudes feel about each other and communicate their attitudes. This is reflected by qualitative research concluding that Dutch participants were "not worried about the existence of differences of opinion per se, but about the broader and more intense expression of these differences. They are also worried about the lack of tolerance and hostility towards those who think differently" (Miltenburg et al., 2022, p. 52). Studies have found that people do not only overestimate attitudinal polarization, but also affective polarization (Moore-Berg et al., 2020). In line with this, Overgaard (2024) found that offering Americans information about low levels of hostility between Republicans and Democrats reduced affective polarization. Similarly, correcting such meta-prejudices by offering information on truthful levels of affective polarization also has the potential to mitigate polarization worry.

The study has several limitations that suggest directions for future research. First, it focused solely on polarization of immigration attitudes. While overestimating polarization appears to be common across various contentious issues (Levendusky & Malhotra, 2016), immigration debates are often particularly hostile, and individuals may be more worried about tone than about attitudinal differences. Thus, the relationship between perceived polarization and polarization worry may be stronger for other topics like climate change, LGBTIQA+ rights, or abortion, where opposing views are prevalent, but the tone may be considered less aggressive. Second, although the operationalization of perceived polarization based on participants' perception of political electorates' attitudes is the prevailing measure in existing literature, it has its limitations. The rather indirect measure might not always reflect respondents' subjective impression of polarization. For example, it may be subject to participants' political knowledge. The four political parties used to measure perceived polarization were selected based on their popularity and salient immigration stances. However, the perceived polarization score may not be an accurate reflection for participants who lacked knowledge about the parties' immigration positions. For example, a score of 0, which means that the participant reported the immigration stances of the more conservative and progressive electorates to be equal, could indicate political ignorance rather than a perception of minimal polarization (even though participants were allowed to skip questions and those who did were excluded from the analysis). Future research could explore alternative operationalizations of perceived polarization, such as examining the perceived distribution of attitudes in the general public (see Van Boven et al., 2012 for an example). Though cognitively demanding, these measures do not require specific knowledge of political parties, which could be useful in political systems with many parties, such as the Netherlands. Third, participants who attended (applied) university were slightly overrepresented in the sample. Although I do not expect this to substantially affect the results, one could argue that higher educated individuals might be more susceptible to information from public opinion research, meaning that the manipulation's effect could be somewhat

weaker in a sample not skewed toward higher education. Finally, the sample was limited to individuals who were, and both their parents were born in the Netherlands. While I have no theoretical reasons to believe this impacted the relationships found, (descendants of) migrants might perceive even greater polarization in immigration attitudes and might be more worried about polarization on this topic, due to personal experiences with prejudice and racism. Future research using more diverse samples should explore this possibility.

Conflict of Interests
The author declares no conflict of interest.

Data Availability
All data and code needed to reproduce the results can be found here: https://osf.io/gpf4w

Supplementary Material
Supplementary material for this article is available online in the format provided by the author (unedited).

References
Abramowitz, A. I., & Saunders, K. L. (2008). Is polarization a myth? *The Journal of Politics*, *70*(2), 542–555. https://doi.org/10.1017/s0022381608080493

Aguinis, H., Beaty, J. C., Boik, R. J., & Pierce, C. A. (2005). Effect size and power in assessing moderating effects of categorical variables using multiple regression: A 30-year review. *Journal of Applied Psychology*, *90*(1), 94–107.

Ahler, D. J. (2014). Self-fulfilling misperceptions of public polarization. *The Journal of Politics*, *76*(3), 607–620.

Amialchuk, A., Ajilore, O., & Egan, K. (2019). The influence of misperceptions about social norms on substance use among school-aged adolescents. *Health Economics*, *28*(6), 736–747. https://doi.org/10.1002/hec.3878

Andre, P., Boneva, T., Chopra, F., & Falk, A. (2024). Misperceived social norms and willingness to act against climate change. *Review of Economics and Statistics*. Advance online publication. https://doi.org/10.1162/rest_a_01468

Baldassarri, D., & Bearman, P. (2007). Dynamics of political polarization. *American Sociological Review*, *72*(5), 784–811.

Banks, A., Calvo, E., Karol, D., & Telhami, S. (2020). #PolarizedFeeds: Three experiments on polarization, framing, and social media. *The International Journal of Press/Politics*, *26*(3), 609–634. https://doi.org/10.1177/1940161220940964

Baumann, F., Lorenz-Spreen, P., Sokolov, I. M., & Starnini, M. (2020). Modeling echo chambers and polarization dynamics in social networks. *Physical Review Letters*, *124*(4). https://doi.org/10.1103/PhysRevLett.124.048301

Blatz, C. W. (2024). A brief intervention teaching false polarization and naïve realism reduces perceived political polarization. *Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement*, *56*(3), 175–186. https://doi.org/10.1037/cbs0000379

Broockman, D. E., & Skovron, C. (2018). Bias in perceptions of public opinion among political elites. *American Political Science Review*, *112*(3), 542–563. https://doi.org/10.1017/s0003055418000011

Carpini, M. X. D., Cook, F. L., & Jacobs, L. R. (2004). Public deliberation, discursive participation, and citizen engagement: A review of the empirical literature. *Annual Review of Political Science*, *7*(1), 315–344. https://doi.org/10.1146/annurev.polisci.7.121003.091630

Cialdini, R. B., Reno, R. R., & Kallgren, C. A. (1990). A focus theory of normative conduct: Recycling the concept of norms to reduce littering in public places. *Journal of Personality and Social Psychology*, *58*(6), 1015–1026. https://doi.org/10.1037/0022-3514.58.6.1015

Clayton, S. (2020). Climate anxiety: Psychological responses to climate change. *Journal of Anxiety Disorders*, *74*, Article 102263. https://doi.org/10.1016/j.janxdis.2020.102263

Cooper, J., Kelly, K. A., & Weaver, K. (2001). Attitudes, norms, and social groups. In M. Hogg (Ed.), *Blackwell handbook of social psychology: Group processes* (pp. 259–282). Blackwell Publishers.

Cowan, S. K., & Baldassarri, D. (2018). "It could turn ugly": Selective disclosure of attitudes in political discussion networks. *Social Networks*, *52*, 1–17. https://doi.org/10.1016/j.socnet.2017.04.002

Dahlgren, P. (2002). In search of the talkative public: Media, deliberative democracy and civic culture. *Javnost-The Public*, *9*(3), 5–25.

Dekker, P. (2022). *Politieke polarisatie in Nederland*. Het Wereldvenster.

Enders, A. M. (2021). Issues versus affect: How do elite and mass polarization compare? *The Journal of Politics*, *83*(4), 1872–1877. https://doi.org/10.1086/715059

Enders, A. M., & Armaly, M. T. (2018). The differential effects of actual and perceived polarization. *Political Behavior*, *41*(3), 815–839. https://doi.org/10.1007/s11109-018-9476-2

Fernbach, P. M., & Van Boven, L. (2022). False polarization: Cognitive mechanisms and potential solutions. *Current Opinion in Psychology*, *43*, 1–6. https://doi.org/10.1016/j.copsyc.2021.06.005

Fiorina, M. P., & Abrams, S. J. (2008). Political polarization in the American public. *Annual Review of Political Science*, *11*(1), 563–588. https://doi.org/10.1146/annurev.polisci.11.053106.153836

Fischer, C. S., & Hout, M. (2006). *Century of difference: How America changed in the last one hundred years*. Russell Sage Foundation.

Gibson, J. L., & Sutherland, J. L. (2023). Keeping your mouth shut: Spiraling self-censorship in the United States. *Political Science Quarterly*, *138*(3), 361–376. https://doi.org/10.1093/psquar/qqad037

Gifford, R., & Nilsson, A. (2014). Personal and social factors that influence pro-environmental concern and behaviour: A review. *International Journal of Psychology*, *49*(3), 141–157. https://doi.org/10.1002/ijop.12034

Graham, M. H., & Svolik, M. W. (2020). Democracy in America? Partisanship, polarization, and the robustness of support for democracy in the United States. *American Political Science Review*, *114*(2), 392–409. https://doi.org/10.1017/s0003055420000052

Herold, M., Joachim, J., Otteni, C., & Vorländer, H. (2023). *Polarization in Europe: A comparative analysis of ten European countries*. Mercator Forum Migration und Demokratie. https://forum-midem.de/midem-study-polarization-in-europe-quantitive-analysis-by-country

Hornsey, M. J., Chapman, C. M., Alvarez, B., Bentley, S., Salvador Casara, B. G., Crimston, C. R., Ionescu, O., Krug, H., Preya Selvanathan, H., Steffens, N. K., & Jetten, J. (2021). To what extent are conspiracy theorists concerned for self versus others? A Covid-19 test case. *European Journal of Social Psychology*, *51*(2), 285–1293. https://doi.org/10.1002/ejsp. 2737

Hossain, M. M., Tasnim, S., Sultana, A., Faizah, F., Mazumder, H., Zou, L., McKyer, E. L. J., Ahmed, H. U., & Ma, P. (2020). Epidemiology of mental health problems in Covid-19: A review. *F1000Research*, *9*, Article 636. https://doi.org/10.12688/f1000research.24457.1

Iyengar, S., Lelkes, Y., Levendusky, M., Malhotra, N., & Westwood, S. J. (2019). The origins and consequences of affective polarization in the United States. *Annual Review of Political Science*, *22*(1), 129–146. https://doi.org/10.1146/annurev-polisci-051117-073034

Iyengar, S., & Westwood, S. J. (2014). Fear and loathing across party lines: New evidence on group polarization. *American Journal of Political Science*, *59*(3), 690–707. https://doi.org/10.1111/ajps.12152

Labrie, J. W., Napper, L. E., & Hummer, J. F. (2014). Normative feedback for parents of college students: Piloting a parent based intervention to correct misperceptions of students' alcohol use and other parents' approval of drinking. *Addictive Behaviors*, *39*(1), 107–113. https://doi.org/10.1016/j.addbeh.2013.08.036

Larimer, M. E., & Neighbors, C. (2003). Normative misperception and the impact of descriptive and injunctive norms on college student gambling. *Psychology of Addictive Behaviors*, *17*(3), 235–243. https://doi.org/10.1037/0893-164X.17.3.235

Lee, A. H. (2022). Social trust in polarized times: How perceptions of political polarization affect Americans' trust in each other. *Political Behavior*, *44*(3), 1533–1554. https://doi.org/10.1007/s11109-022-09787-1

Lees, J., & Cikara, M. (2021). Understanding and combating misperceived polarization. *Philosophical Transactions of the Royal Society*, *376*(1822), Article 20200143. https://doi.org/10.1098/rstb.2020.0143

Lelkes, Y. (2016). Mass polarization: Manifestations and measurements. *Public Opinion Quarterly*, *80*(S1), 392–410. https://doi.org/10.1093/poq/nfw005

Levendusky, M. S. (2018). Americans, not partisans: Can priming American national identity reduce affective polarization? *The Journal of Politics*, *80*(1), 59–70. https://doi.org/10.1086/693987

Levendusky, M. S., & Malhotra, N. (2016). (Mis)perceptions of partisan polarization in the American public. *Public Opinion Quarterly*, *80*(S1), 378–391. https://doi.org/10.1093/poq/nfv045

Lintonen, T. P., & Konu, A. I. (2004). The misperceived social norm of drunkenness among early adolescents in Finland. *Health Education Research*, *19*(1), 64–70. https://doi.org/10.1093/her/cyg010

Matthes, J., Knoll, J., & von Sikorski, C. (2017). The "spiral of silence" revisited: A meta-analysis on the relationship between perceptions of opinion support and political opinion expression. *Communication Research*, *45*(1), 3–33. https://doi.org/10.1177/0093650217745429

Miltenburg, E., Geurkink, B., Tunderman, S., Beekers, D., & Ridder, J. d. (2022). *Burgerperspectieven 2022: 2*. Sociaal en Cultureel Planbureau. https://www.scp.nl/publicaties/publicaties/2022/12/29/continu-onderzoek-burgerperspectieven---bericht-2-2022

Moore-Berg, S. L., Ankori-Karlinsky, L. O., Hameiri, B., & Bruneau, E. (2020). Exaggerated meta-perceptions predict intergroup hostility between American political partisans. *Proceedings of the National Academy of Sciences*, *117*(26), 14864–14872. https://doi.org/10.1073/pnas.2001263117

Moral, M. (2017). The bipolar voter: On the effects of actual and perceived party polarization on voter turnout in European multiparty democracies. *Political Behavior*, *39*(4), 935–965. https://doi.org/10.1007/s11109-016-9386-0

Muis, Q. (2024). *"Who are those people?": Causes and consequences of polarization in the schooled society*. Open Press Tilburg University. https://doi.org/https://doi.org/10.56675/tsb.27333853

Myers, D. G., & Lamm, H. (1976). The group polarization phenomenon. *Psychological Bulletin*, *83*(4), 602–627.

Nayak, S. S., Fraser, T., Panagopoulos, C., Aldrich, D. P., & Kim, D. (2021). Is divisive politics making Americans sick? Associations of perceived partisan polarization with physical and mental health outcomes among adults in the United States. *Social Science & Medicine*, *284*, Article 113976. https://doi.org/10.1016/j.socscimed.2021.113976

Niemiec, R. M., Champine, V., Vaske, J. J., & Mertens, A. (2020). Does the impact of norms vary by type of norm and type of conservation behavior? A meta-analysis. *Society & Natural Resources*, *33*(8), 1024–1040. https://doi.org/10.1080/08941920.2020.1729912

Noelle-Neumann, E. (1974). The spiral of silence a theory of public opinion. *Journal of Communication*, *24*(2), 43–51.

NOS. (2022). *Nederlanders ergeren zich aan polarisatie, omringen zich vooral met gelijkgestemden.* https://nos.nl/artikel/2458038-nederlanders-ergeren-zich-aan-polarisatie-omringen-zich-vooral-met-gelijkgestemden

Orhan, Y. E. (2022). The relationship between affective polarization and democratic backsliding: Comparative evidence. *Democratization*, *29*(4), 714–735. https://doi.org/10.1080/13510347.2021.2008912

Overgaard, C. S. B. (2024). Perceiving affective polarization in the United States: How social media shape meta-perceptions and affective polarization. *Social Media + Society*, *10*(1). https://doi.org/10.1177/20563051241232662

Perkins, H. W., Craig, D. W., & Perkins, J. M. (2011). Using social norms to reduce bullying. *Group Processes & Intergroup Relations*, *14*(5), 703–722. https://doi.org/10.1177/1368430210398004

Peters, U. (2021). How (many) descriptive claims about political polarization exacerbate polarization. *Journal of Social and Political Psychology*, *9*(1), 24–36. https://doi.org/10.5964/jspp. 5543

Rand, D. G., & Yoeli, E. (2024). Descriptive norms can "backfire" in hyper-polarized contexts. *PNAS Nexus*, *3*(10). https://doi.org/10.1093/pnasnexus/pgae303

Reiljan, A. (2019). *The most polarising issues in the European Union party system.* EU&I. https://euandi.eu/en/blog/knowlegde-bites/the-most-polarising-issues-in-the-european-union-party-system.html

Robison, J., & Mullinix, K. J. (2015). Elite polarization and public opinion: How polarization is communicated and its effects. *Political Communication*, *33*(2), 261–282. https://doi.org/10.1080/10584609.2015.1055526

Roblain, A., & Green, E. G. T. (2021). From perceived polarization of immigration attitudes to collective action. *International Journal of Intercultural Relations*, *80*, 112–120. https://doi.org/10.1016/j.ijintrel.2020.11.009

Ros, R. (2023). *Polarisatiepaniek.* De Groene Amsterdammer. https://www.groene.nl/artikel/polarisatiepaniek

SIRE. (2023). *Verlies elkaar niet: als polarisatie dichtbij komt—Infographic SIRE campagne.* https://polarisatieverlieselkaarniet.sire.nl/wp-content/uploads/2023/01/Sire_Polarisatie_Infographic.pdf

Skelley, G., & Fuong, H. (2022). *3 in 10 Americans named political polarization as a top issue facing the country.* FiveThirtyEight. https://fivethirtyeight.com/features/3-in-10-americans-named-political-polarization-as-a-top-issue-facing-the-country

Smith, K. B. (2022). Politics is making us sick: The negative impact of political engagement on public health during the Trump administration. *PLoS One*, *17*(1), Article 0262022. https://doi.org/10.1371/journal.pone.0262022

Smith, K. B., Hibbing, M. V., & Hibbing, J. R. (2019). Friends, relatives, sanity, and health: The costs of politics. *PLoS One*, *14*(9), Article 0221870. https://doi.org/10.1371/journal.pone.0221870

Van Boven, L., Judd, C. M., & Sherman, D. K. (2012). Political polarization projection: Social projection of partisan attitude extremity and attitudinal processes. *Journal of Personality and Social Psychology*, *103*(1), 84–100. https://doi.org/10.1037/a0028145

Voogd, R., Jacobs, K., Lubbers, M., & Spierings, N. (2024). *De verkiezingen van 2023. Van onderstroom naar doorbraak: Onvrede en migratie.* SKON.

Webster, S. W., & Abramowitz, A. I. (2017). The ideological foundations of affective polarization in the US electorate. *American Politics Research*, *45*(4), 621–647. https://doi.org/10.1177/1532673x17703132

Wenzel, M. (2005). Misperceptions of social norms about tax compliance: From theory to intervention. *Journal of Economic Psychology*, *26*(6), 862–883. https://doi.org/10.1016/j.joep. 2005.02.002

Westfall, J., Van Boven, L., Chambers, J. R., & Judd, C. M. (2015). Perceiving political polarization in the United States: Party identity strength and attitude extremity exacerbate the perceived partisan divide. *Perspectives on Psychological Science*, *10*(2), 145–158. https://doi.org/10.1177/1745691615569849

Wilson, A. E., Parker, V. A., & Feinberg, M. (2020). Polarization in the contemporary political and media landscape. *Current Opinion in Behavioral Sciences*, *34*, 223–228. https://doi.org/10.1016/j.cobeha.2020.07.005

World Economic Forum. (2024). *The global risks report 2024: 19th edition* (Vol. 19). World Economic Forum.

Wuestenenk, N., Nijs, T., Stark, T. H., van Tubergen, F., & Ellemers, N. (2025). The interplay of misperceptions and willingness to share opinions in full classroom networks: The case of opinions towards homosexuality. *Social Networks*, *80*, 25–35. https://doi.org/10.1016/j.socnet.2024.08.003

Yang, J., Rojas, H., Wojcieszak, M., Aalberg, T., Coen, S., Curran, J., Hayashi, K., Iyengar, S., Jones, P. K., Mazzoleni, G., Papathanassopoulos, S., Rhee, J. W., Rowe, D., Soroka, S., & Tiffen, R. (2016). Why are "others" so polarized? Perceived political polarization and media use in 10 countries. *Journal of Computer-Mediated Communication*, *21*(5), 349–367. https://doi.org/10.1111/jcc4.12166

About the Author

Tom Nijs is a postdoctoral researcher at the European Research Center on Migration and Ethnic Relations (ERCOMER) and the Interuniversity Center for Social Science Theory and Methodology (ICS). His research interests include interethnic attitudes, ethnic identification, intergroup contact, polarization, ownership beliefs, and radical right-wing voting.

ARTICLE

How Descriptive Norms and Peer Attitudes Shape Interethnic Dating Among Adolescents in Dutch Schools

Eva Jaspers [1], Pascale I. van Zantvliet [2,3], and Jan-Willem Simons [1]

[1] Department of Sociology, Utrecht University, The Netherlands
[2] Department of Education and Pedagogy, Utrecht University, The Netherlands
[3] Utrecht Center for Research and Development of Health Professions Education, University Medical Center Utrecht, The Netherlands

Correspondence: Eva Jaspers (e.jaspers@uu.nl)

Abstract

This study examines the role of school-based environmental factors in shaping interethnic dating patterns among adolescents in the Netherlands. Given that adolescence is a formative period for social behaviour, early interethnic dating experiences may foster openness to interethnic marriage later in life. We focus on how general peer approval of the outgroup—from both the ingroup and outgroup perspectives—and descriptive norms (i.e., the prevalence of interethnic relationships) influence adolescents' partner choices. Using longitudinal data from the Dutch part of the Children of Immigrants Longitudinal Survey in Four European Countries (CILS4EU), we test whether ingroup attitudes influence interethnic dating via personal preferences for interethnic dating, and whether descriptive norms and outgroup attitudes moderate the relationship between personal preferences and dating. Results support the hypothesis that more positive ingroup attitudes are indirectly associated with an increased likelihood of interethnic dating via personal preferences. Descriptive norms furthermore moderate the effect of personal preferences, strengthening the relationship under supportive normative conditions. In contrast, we find no evidence that outgroup attitudes towards the ingroup play a moderating role. Ethnic minority adolescents and older students are more likely to date someone from an ethnic outgroup, while students with more highly educated parents and Christian students (vs. non-religious peers) are less likely to do so. Future research should explore additional sources of normative influence, such as (social) media and adolescent clubs, to better understand the broader forces shaping interethnic relationships.

Keywords

adolescents; descriptive norms; intergroup attitudes; interethnic dating; schools

---

1. Introduction

The increasing diversity in Europe, driven by migration and low native fertility rates (Coleman, 2006), has created greater opportunities for individuals to interact across ethnic lines. This enhanced interaction has the potential to foster more interethnic marriages, which are significant markers of social integration and mutual acceptance between groups. Interethnic marriage is widely regarded as a crucial indicator of integration, as it reflects the degree to which individuals overcome group boundaries. Accepting an out-group member as a spouse suggests that group differences are no longer seen as barriers to long-term partnerships (Alba & Nee, 2003). It signals a reduction in prejudice and the successful blending of cultural differences within a society.

Adolescence is a particularly formative period for social and romantic relations, significantly influencing partner choice in adulthood. Early interethnic dating experiences may shape attitudes towards out-group members and increase the likelihood of intermarriage later in life (van Zantvliet et al., 2014a). Studies show that adolescent romantic experiences often persist into adulthood (Madsen & Collins, 2011; Raley et al., 2007), and those with early interethnic relationships may display a greater openness to intermarriage (King & Harris, 2007). The cues on the appropriateness of interethnic contact within schools encompass both descriptive norms—characterized by the prevalence of certain behaviors in a given social context—and general attitudes of the environment—characterized by the perception of what most people approve or disapprove of (Cialdini et al., 1991). Together, these cues reflect the social acceptability of interethnic dating (Bourgeois & Leary, 2001; Kalmijn, 1998).

Norms in relevant dating settings are pivotal for understanding behavior (Cialdini et al., 1991; Durkheim, 1951, yet research on interethnic dating often neglects the influence of local norms, which we hypothesize may play a pivotal role in shaping romantic partner choices. For adolescents, schools serve as crucial environments for forming romantic connections and friendships. Normative influences in schools are particularly significant because adolescents spend much of their time in these settings, and schools provide one of the primary meeting grounds for romantic partners (Kiesner et al., 2004; Mollenhorst et al., 2008). Schools are spaces where norms regarding interethnic relationships can both be observed and transmitted (van Vemde et al., 2021). Descriptive norms such as interethnic dating patterns within schools reveal the openness of its students to interethnic interactions, thereby encouraging or discouraging interethnic dating (Vaquera & Kao, 2008).

Schoolmates' explicit interethnic attitudes—students' consciously reported evaluations of ethnic groups—influence the social dynamics within ethnically diverse secondary school classrooms (van Vemde et al., 2023). Explicit interethnic attitudes are linked to deliberate and consciously monitored behaviors, including verbal and socially strategic interactions (Azjen et al., 2018). Given that students' social experiences in the school can be shaped by the evaluations of others about their ethnic group, and that these attitudes reflect socially relevant judgments, it is plausible that school interethnic attitudes influence the development of interethnic relationships (Bayram Özdemir et al., 2018; König et al., 2022; Thijs et al., 2014).

In this study, we investigate how the school environment influences interethnic dating between native and ethnic minority adolescents (with Turkish, Moroccan, Surinamese, and Antillean heritage) in the Netherlands. Using the Children of Immigrants Longitudinal Survey in Four European Countries (CILS4EU; Kalter et al., 2014), we analyze how interethnic attitudes and descriptive norms within schools shape partner choices. The longitudinal design allows us to examine whether school-based attitudes and norms on interethnic

relationships indeed influence dating partner choices, while the inclusion of data from several classes per school provides us with a reliable measure of the interethnic attitudes and interethnic relationships within the school context.

2. Theory

2.1. The Basic Model: Preferences, Opportunities, and Third Parties

A review of the sociological literature on intermarriage (Kalmijn, 1998) has pointed to three factors that influence the choice of a partner: (a) personal and (b) third-party preferences for certain characteristics in a partner, as well as (c) opportunities of the marriage market in which candidates search for a partner. We apply this framework and specify the normative component to explain the choice of a dating partner by native and ethnic minority adolescents in the Netherlands (see also van Zantvliet et al., 2014a; Weißmann & Maddox, 2016).

Partner preferences are first held by individuals. In general, individuals prefer to interact and establish relationships with similar others (Feliciano & Hijara, 2024; Kalmijn, 1998; Lewis, 2016; McPherson et al., 2001). Cultural similarity enhances personal attraction because it confirms one's worldview, provides opportunities for joint activities, and facilitates mutual understanding (Kalmijn, 1998). This preference for endogamy is already expressed in adolescence (Blackwell & Lichter, 2004; Herman & Campbell, 2012; Weißmann & Maddox, 2016).

Preferences for endogamy are also held by third parties, such as friends and family (Huijnk & Liefbroer, 2012; Huijnk et al., 2012; Lundquist et al., 2024). Exogamy may threaten affiliation with the own group (Finnäs & O'Leary, 2003; Qian, 2004), solidarity with the own group (Kalmijn, 1998), and the family reputation (Munniksma et al., 2012). Endogamy, on the other hand, may ensure the intergenerational transmission of the group's culture (Kalmijn, 2015; Xie & Goyette, 1997), and ease communication and communal activity (McPherson et al., 2001).

Besides personal and third-party preferences, the formation of romantic relations depends on social structures (Blau, 1977). The realization of preferences is, thus, dependent on the meeting opportunities of the social context: Relationships are only possible with those people one has the opportunity to meet (Blau, 1977; Verbrugge, 1977), and relationships with members of certain groups are more likely if one's social contexts consist of relatively more members of these groups (de Vroome et al., 2014; Harris & Ono, 2005; Kalmijn & van Tubergen, 2010; van Tubergen & Maas, 2007). In adolescence, the school is a social context that is particularly relevant for finding new friends and partners (Kiesner et al., 2004), because of the large amount of time that is spent in this context, and because adolescents have limited mobility. The opportunities for intergroup contact that are offered by this social context have been found to affect adolescents' partner choices (Fujino, 1997; Shibazaki & Brennan, 1998; Wang & Kao, 2007), hence we control for the opportunities for interethnic dating in our analytical approach.

The model of preferences and opportunities for romantic partner choice has proven very helpful in explaining interethnic partner choices and it is therefore adopted as a basic model in this study. We extend the model by refining and examining the specific influences that third parties exert on interethnic dating.

Specifically, we investigate how general attitudes regarding the outgroup—from both the ingroup and the outgroup—as well as descriptive norms of interethnic relationships, shape interethnic dating patterns. For clarity, we hypothesize on a situation in which there are only two groups, of which both members view all individuals in their own group as the ingroup, and all individuals in the other group as the outgroup. We contend that local cues for appropriate behavior indirectly influence interethnic dating in three ways. First, the outgroup attitudes of the ingroup inform individuals' personal preferences for interethnic relationships, which in turn affect their likelihood of engaging in interethnic dating (Kalmijn, 1998). Second, the outgroup attitudes of the outgroup reflect outgroup members' willingness to reciprocate interethnic relationship attempts, resulting in the extent to which individuals can translate personal interethnic dating preferences, if any, into actual interethnic relationships (beyond mere opportunity). Third, descriptive norms indicate to what extent interethnic dating would be considered appropriate or desirable behavior by peers, and whether adolescents with interethnic dating partners would face peer penalties. We argue that these descriptive norms influence the extent to which individuals would be willing to act on their own preference for interethnic dating. See Figure 1 for a graphic overview of the hypothesized effects.
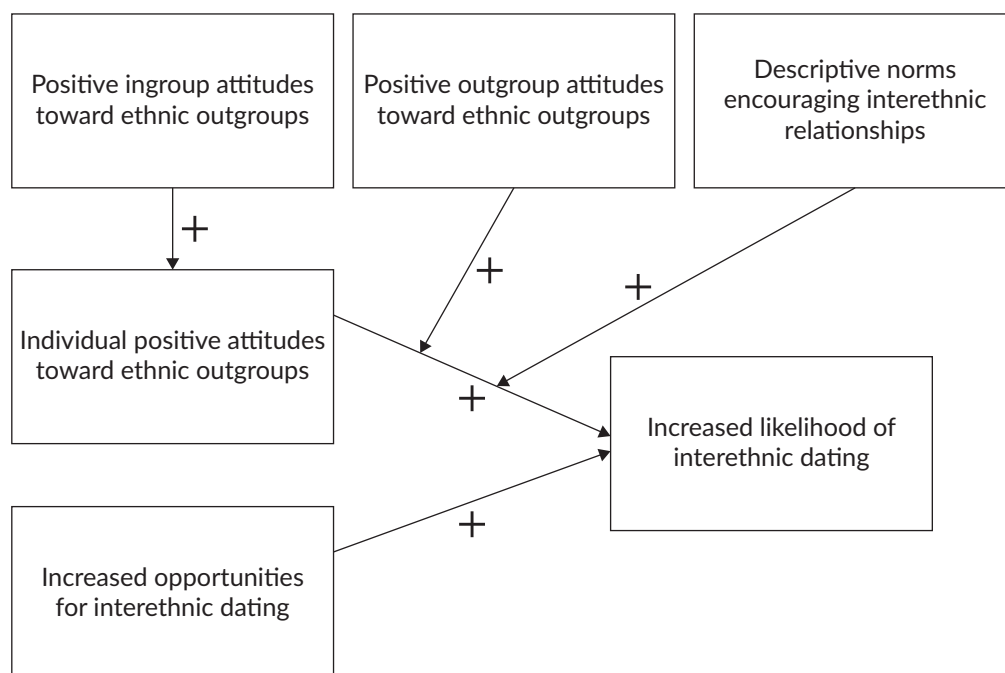
Figure 1. Extended theoretical model of preferences, opportunities, and third-party influences on interethnic dating.

2.2. The Influence of Ingroup Schoolmates' Interethnic Attitudes on Interethnic Dating

As outlined in the previous section, personal and third-party attitudes play an important role in shaping interethnic behaviors like interethnic dating (Kalmijn, 1998). The influence of third-party attitudes on interethnic behaviors is typically theorized to operate through their effect on personal preferences, by shaping the extent to which individuals conform to the general attitudes regarding ethnic outgroups (Kalmijn, 1998).

Generalized attitudes reflect what is commonly approved of within a social group and function as a mechanism of social regulation. Approval of an ethnic outgroup by relevant ingroup members—particularly peers or other significant referents—can influence individuals' interethnic relational preferences and behaviors through several interrelated sociological mechanisms. First, normative influence plays a central role: When ingroup members express positive attitudes towards an outgroup, they signal socially acceptable boundaries for interaction, thereby reducing normative constraints on forming interethnic ties (Rivas-Drake et al., 2017). This social approval can foster a perception that intergroup relationships are acceptable, or even desirable, within the peer group, thereby encouraging individual engagement in such relationships. Second, individuals often look to salient peer groups to inform their attitudes and behaviors (Merton, 1941). If peers demonstrate acceptance or preference for interethnic relations, individuals may internalize these orientations, adjusting their own relational preferences accordingly. Finally, social network dynamics may play a crucial mediating role. Positive group-level attitudes can lead to increased intergroup contact opportunities (Moody, 2001), lowering the social cost of crossing group boundaries and fostering mutual trust and familiarity (Pettigrew, 1998). This is particularly relevant in settings such as schools, where peer approval is closely tied to social capital and reputation (Cillessen & Rose, 2005). Taken together, these mechanisms suggest that positive ingroup attitudes towards an ethnic outgroup create a social context in which interethnic relationships are more likely to be preferred, pursued, and successfully maintained.

In the context of interethnic dating, ingroup attitudes towards the outgroup are thus expected to shape interethnic dating preferences. When these attitudes are positive towards the outgroup, individuals anticipate social or internal approval by their ingroup, leading them to develop more favorable personal preferences for interethnic dating, and consequently, a greater likelihood of engaging in such relationships (Grusec & Hastings, 2014). Based on this theoretical framework, we formulate the following hypothesis:

> Hypothesis 1: Positive ingroup attitudes towards the outgroup will result in an increased personal preference for interethnic dating, which will, in turn, lead to a greater likelihood of engaging in interethnic dating.

2.3. The Influence of Outgroup Attitudes on the Realization of Preferences

In addition to ingroup attitudes, we consider the role of outgroup attitudes towards the outgroup (i.e., the individual's ingroup) in shaping interethnic dating behavior of the individual. While ingroup interethnic attitudes directly influence personal preferences for interethnic dating, we theorize that outgroup attitudes towards the individual's ingroup moderate the extent to which these preferences translate into actual interethnic dating behavior. Outgroup attitudes regulate the social acceptability of interethnic partners within the outgroup. We postulate that this shapes the degree to which individuals can act on their personal preferences for interethnic dating, a distinction that has largely been overlooked in discussions of third-party influence on interethnic romantic relationships (Kalmijn, 1998).

Theoretically, outgroup attitudes function similarly to ingroup attitudes but operate within the ethnic outgroup. Rather than shaping personal preferences directly, they moderate the extent to which individuals can act on their existing preferences for interethnic dating. Even if an individual holds a strong preference for interethnic dating and has opportunities to engage in such relationships because of enough outgroup members—and the ingroup generally holds positive attitudes towards such relationships—the extent to

which they can pursue them depends on the interethnic attitudes of the ethnic outgroup (Tropp & Pettigrew, 2005). When outgroup interethnic attitudes are positive towards the ingroup, individuals who are positive about an outgroup member may be more likely to act on their preference, as they perceive increased openness from outgroup members. Furthermore, outgroup members may be more likely to reciprocate such dating attempts. Conversely, when outgroup attitudes disapprove of the ingroup, the influence of personal preference on interethnic dating behavior weakens, as individuals may face rejection from outgroup members. On that basis, we formulate the following hypothesis:

> Hypothesis 2: Outgroup interethnic attitudes that are positive towards the ingroup will moderate the relationship between personal preference for interethnic dating and the likelihood of engaging in interethnic dating, such that this relationship will be stronger when outgroup attitudes are more positive towards the ingroup.

2.4. The Influence of Descriptive Norms on Interethnic Dating

Descriptive norms indicate what behaviors are prevalent within a given social context and function as social heuristics, guiding behavior by signaling which actions provide the most certainty for navigating a social group effectively (Cialdini et al., 1991). Research has shown that individuals rely on descriptive norms to guide behavior, particularly in ambiguous or uncertain social situations such as the adolescent dating market (Cialdini et al., 1991).

In the context of interethnic dating, descriptive norms provide individuals with cues about the prevalence of interethnic relationships and, consequently, their feasibility. When interethnic contact is widely observed, individuals infer that such behavior is socially acceptable and that outgroup members are receptive to such relationships, reinforcing the perception that interethnic dating is a viable option (Tropp & Pettigrew, 2005). Conversely, when interethnic contact is rare or unseen, individuals may perceive it as socially uncertain or undesirable, anticipating a higher likelihood of rejection from the ethnic outgroup (Herman & Campbell, 2012).

Descriptive norms are hypothesized to moderate the extent to which personal preferences translate into actual dating behavior by shaping perceptions of feasibility. When interethnic dating is prevalent and normalized, individuals who are positive towards outgroup members and prefer an interethnic relationship are more likely to act on their preferences, as they perceive a lower risk of rejection from the ethnic outgroup (Herman & Campbell, 2012). Since dating at this age is still rare, the relationships that are observed—even if it is just one—send out a strong signal about the social acceptance and receptiveness of the other group. Conversely, when interethnic dating is absent, individuals may hesitate to pursue such relationships due to the lack of clear social cues which indicates that outgroup members are unlikely to reciprocate interest. Based on this reasoning, we formulate the following hypothesis:

> Hypothesis 3: Descriptive norms encouraging interethnic contact will moderate the relationship between personal preference for interethnic dating and the likelihood of engaging in interethnic dating, such that this relationship will be stronger when descriptive norms signal that interethnic dating is more common.

3. Data and Methods

3.1. Data

We use the first three waves of the CILS4EU (Kalter et al., 2016a, 2016b, 2016c, 2016d) to test our hypotheses. This panel followed adolescents in England, Germany, the Netherlands, and Sweden annually since 2010, when these adolescents were around 15 years old. Using comparable designs and measures, the survey aims to answer key questions about the structural, cultural, and social integration of the children of immigrants in these countries (Kalter et al., 2014). In this study, we restrict our analyses to the Netherlands.

The survey used a stratified three-stage sample design to allow for oversampling schools with a high proportion of pupils with an immigrant background. In the first stage, schools were sampled from a comprehensive national school list with probability proportional to size. The school response rate was 34.9%. For each non-participating school, a school comparable in educational track and immigrant proportion was sampled. This replacement strategy increased coverage of the school sample to 91.7%. In the second stage, within each participating school, at least two classes of the 3rd grade (age about 14–15) were sampled. In the third stage, all pupils enrolled in these classes were sampled. The sampled pupils for whom passive parental consent was granted completed an extensive questionnaire at school. The pupil response rate was 91.1%. In wave 2, respondents who were still in school were again surveyed at school. Respondents who were no longer in school were surveyed via an online questionnaire. In wave 3, respondents were surveyed via online questionnaires. Respondents who were still in school were approached via their school if the school agreed to participate one more time. Conditional on participation in waves 1 or 2 (during which the necessary contact information was collected), response rates for wave 3 are 57.9%. In addition, a parental survey was conducted in wave 1 (and wave 2, for those parents who had not participated in wave 1), for which the response rate was 74.7%.

3.2. Observed Sample

Our observed sample consists of 511 respondents, including 430 ethnic majority members and 81 ethnic minority members. The observed sample was obtained through a series of selection steps. First, we included only those who participated in wave 1 ($n = 4{,}636$) and who self-identified as part of either the ethnic majority or one of the minority groups of interest ($n = 3{,}445$). Second, we limited the sample to respondents who attended a school with at least one student from both the majority and minority group ($n = 2{,}995$). We furthermore excluded three schools where only one student from either group was present who had missing data on the individual attitudes variable ($n = 2{,}883$). Additionally, respondents older than 20 years were excluded ($n = 2{,}776$). Finally, we selected only those respondents who were in a (intra- or interethnic) romantic relationship at either wave 2 or wave 3 and with complete data on the variables of interest, leading to a final sample size of 511. Exploratory Heckman (1979) selection models subsequently indicated that adolescents who were dating did not differ substantially from non-dating adolescents on key characteristics relevant to partner choice, mitigating concerns about selection bias (results available upon request).

3.3. Imputed Sample Sensitivity Analysis

As we lost a substantial number of cases in the observed sample due to the presence of missing data, particularly on the dating status variable in wave 2 (~23%) and wave 3 (~43%), a considerably reduced number of cases was available for analysis. To mitigate potential estimation bias due to data not missing at random and to enhance statistical power, we supplemented our complete-case analysis with a sample obtained by performing cutting-edge multilevel multiple imputation.

Multilevel multiple imputation was conducted using the *mice* package in R (van Buuren, 2018; van Buuren & Groothuis-Oudshoorn, 2011), employing a two-level imputation model to account for the hierarchical structure of the data (i.e., students nested within schools). The imputation was performed on raw level 1 variables, which were subsequently used to derive the key dependent and independent variables at levels 1 and 2 for the analyses.

To inform the imputation model, we selected 110 auxiliary variables based on three criteria: (a) less than 50% missing data, (b) its theoretical relevance with respect to or an absolute correlation of at least .30 with one or more of the dependent and independent variables, and (c) no substantial computational issues during model fitting. A custom visit sequence was used, prioritizing the imputation of auxiliary variables with the least missingness. Each auxiliary variable was sequentially imputed using the fully observed independent variables—specifically, ethnic background and a "mixed parents" indicator (defined as having one immigrant and one native parent)—as well as any previously imputed auxiliary variables. Once all auxiliary variables were imputed, they were used in combination with the fully observed independent variables to impute the remaining dependent and independent variables, without applying a visit sequence.

The school identifier was specified as the clustering variable in the predictor matrix, while the method "2l.pmm" was used for imputing all variables. A total of 40 imputed datasets were generated using 20 iterations each. Diagnostic checks—including trace plots for convergence and density plots to assess imputation plausibility—indicated that the imputations performed sufficiently.

3.4. Measures

3.4.1. Dependent Variable

Interethnic dating was a binary measure indicating whether an adolescent was in an intra- (0) or interethnic romantic relationship (1) in either wave 2 or 3. In these waves, respondents were asked: "Do you currently have a boyfriend or girlfriend?" Those who answered affirmatively were then asked about their partner's background, with response options including "Dutch," "Antillean," "Surinamese," "Turkish," "Moroccan," and "Other background." The definition of background was left to the respondent's interpretation, which could refer to ethnic or national origin of their partner, their partner's parents, or their (great-)grandparents. Based on the respondent's own national origin (see Section 3.4.4), relationships were classified as intra-ethnic (0) if the partner belonged to the same background and interethnic (1) if the partner was from a different background. The "Other background" category was excluded from this classification. We now turn to a discussion of the independent variables, all of which were measured at wave 1.

3.4.2. Independent Variables

Respondents' *individual attitudes towards the ethnic outgroup* was operationalized as their attitude towards the ethnic outgroup, measured using a thermometer question. This question asked respondents to indicate their feelings towards the ethnic outgroup on a scale from 0 to 100, in increments of 10, where 0 represented *very negative* feelings, 50 was *neutral*, and 100 was *very positive*. A higher score on this variable thus indicates a more positive individual attitude towards the outgroup. For native Dutch individuals, the score was calculated as the mean rating across the thermometer questions for the four ethnic minority groups (Turkish, Moroccan, Surinamese, Antillean; $\alpha_{obs} = 0.90$, $\alpha_{imp} = .90$). For each ethnic minority group, the score reflected their attitude towards native Dutch people.

For each respondent, the *ingroup attitude towards the ethnic outgroup* was operationalized as the average attitude of their ethnic ingroup towards the ethnic outgroup within their respective school. For the majority (native Dutch) respondents, the average ingroup attitude was thus calculated as the average attitude of their native Dutch peers within the same school towards ethnic minority groups. For ethnic minority respondents, the average ingroup attitude reflected the average attitude of members of their ethnic ingroup within a school towards the native Dutch population. Conversely, the *outgroup attitude towards the ethnic ingroup* reflected the average view of the respondent's ethnic outgroup towards their ingroup: For native Dutch respondents, this was average attitude of ethnic minority peers towards natives; for ethnic minority respondents, it was the average attitude of native Dutch peers towards minorities.

The variable *descriptive interethnic relationship norms* was measured as the total number of romantic relationships between ethnic majority and minority individuals within the sampled classes in a given school. As it reflects the overall prevalence of interethnic relationships, this measure is identical for both majority and minority individuals within the same school. A higher value indicates a greater occurrence of interethnic relationships within a school. We also include the proportion as an alternative to the absolute value, because it accounts for differences in school size and relationship prevalence.

3.4.3. Control Variables: Personal Preferences, Third-Party Preferences, and Opportunities

Besides the main variable of interest, our analyses included variables that have been associated with three factors from the basic model of partner choice: *personal preferences*, *third-party preferences*, and *opportunities*.

To account for personal preferences that are associated with a higher likelihood of interethnic dating (van Zantvliet et al., 2014a), we controlled for traditional gender roles and conservative family values. *Traditional gender roles* were assessed by asking respondents whether the following tasks should be done mostly by the man, mostly by the woman, or by both about the same: (a) taking care of the children, (b) cooking, (c) earning money, and (d) cleaning the house. We constructed a scale by counting the number of times the adolescent endorsed the traditional task division, where taking care of the children, cooking, and cleaning should be done mostly by the woman, and earning money should be done mostly by the man (Loevinger's $H_{obs} = 0.48$; Loevinger's $H_{imp} = 0.48$). *Conservative family values* were measured based on respondents' approval of the following behaviors: (a) living together as a couple without being married, (b) divorce, (c) abortion, and (d) homosexuality. Each behavior was rated on a 4-point scale (1 = *always OK*, 2 = *often OK*, 3 = *sometimes OK*, 4 = *never OK*). We calculated the final measure as the average score across items ($\alpha_{obs} = 0.71$; $\alpha_{imp} = 0.70$).

Second, we controlled for parental ethnic background, parental education, and religious affiliation, as these factors have been found to be associated with parental pressure to date within one's own ethnic group (van Zantvliet et al., 2014b) and with their children's interethnic partner choice (van Zantvliet et al., 2014a). For these measures, we used information provided by the parent in the parental survey or, if missing (or if referring to a non-biological or non-adoptive parent), the respective information provided by the child. The dummy variable *mixed parents* indicated that at least one parent had an immigrant background and the other had a native background. The dummy variable *educated parents* indicated that at least one biological or adoptive parent had completed upper secondary school. Parents participating in the parental survey specified the highest education completed by themselves and their partner on a scale with the following answer categories: 1 = *no school leaving certificate*, 2 = *degree below upper secondary school*, 3 = *degree from upper secondary school*, and 4 = *university degree*. Adolescents separately reported whether their biological mother and biological father had completed primary school, upper secondary school, or university. *Religious affiliation* was measured in wave 1 by asking parents about to which religion they belonged, with response options including several prelisted religions and an open-ended category. In our model, we distinguish between "no religion," "Christianity," "Islam," and "other religion" (which combines all other reported affiliations into one category).

Third, the choice of a partner is constrained by the opportunities available in the local social environment (Kalmijn & Flap, 2001). To control for the opportunity structure within schools, we included the proportion of outgroup members (among the opposite sex) within the school as a measure of the availability of potential partners from different ethnic backgrounds.

3.4.4. Control Variables: Demographic Characteristics

We finally controlled for three demographic variables: age, gender, and ethnic majority/minority status, as these factors have been shown to be correlated with interethnic contact (Martinović, 2013) and for dating behaviors (Connolly et al., 2004, 2013; van Zantvliet et al., 2014a). *Age* was measured as a continuous variable, recorded in months at wave 1. *Gender* was included as a dummy variable, with respondents coded as a boy (0) or a girl (1). *Ethnic majority/minority status* was also included as a dummy variable, determined based on the respondent's national origin, using information from wave 1 supplemented with data from waves 2 and 3 to reduce missing values. National origin was coded following the strategy used by Dollmann et al. (2014), which relied on information about the respondent's country of birth as well as the birthplaces of their biological parents and grandparents. Respondents were categorized as the native Dutch ethnic majority (0) if none of their ancestors were foreign-born. If one or more ancestors were foreign-born and from the same country, national origin was assigned to that country. If ancestors were born in different countries or if information about their birthplaces was missing, national origin was determined based on the respondent's own country of birth. Based on this classification, Antillean, Surinamese, Turkish, and Moroccan respondents were assigned to the ethnic minority category (1).

3.5. Descriptives

Table 1 presents the mean, standard deviation (for non-categorical variables), and range for the dependent, independent and control variables in the observed sample. The Supplementary File presents descriptive statistics for the observed sample by ethnic group, along with all descriptives (i.e., across the entire sample

Table 1. Descriptives of dependent and independent variables for the observed sample.

Variable	N	M (SD)	Min-Max
Interethnic dating	511	0.14	0–1
Individual attitudes	511	57.69 (21.69)	0–100
Ingroup attitudes	511	57.16 (9.81)	33.72–95.00
Outgroup attitudes	511	71.74 (12.13)	41.83–100.00
Descriptive norm (abs.)	511	0.95 (1.07)	0–4
Descriptive norm (rel.)	511	2.23 (2.78)	0.00–15.39
Trad. gender roles	511	1.52 (1.31)	0–4
Cons. family values	511	1.28 (0.66)	0–3
Mixed parents	511	0.07	0–1
Educated parents	511	0.96	0–1
Religious affiliation	511		0–3
No religion (0)	315		
Christianity (1)	142		
Islam (2)	30		
Other religion (3)	24		
Proportion outgroup	511	10.51 (11.98)	0.00–65.00
Age (in months)	511	203.9 (7.77)	157–231
Gender (1 = girl)	511	0.66	0–1
Ethnic majority/minority status (1 = min.)	511	0.16	0–1

and by ethnic group) for the imputed sample. We note that the descriptives have not been weighted for the sampling design and may therefore not be fully representative of the population of adolescents who were in the target grade in 2010.

In the observed sample, 14% of adolescents reported being in an interethnic romantic relationship. Individual attitudes towards the ethnic outgroup were slightly more positive than neutral, with a mean score of 57.69 ($SD = 21.69$). The average ingroup attitude towards the outgroup was 57.16 ($SD = 9.81$), while the average outgroup attitude towards the ingroup was higher—71.74 ($SD = 12.13$). The average number of interethnic relationships per school, as an absolute measure of the descriptive norm, was 0.95 ($SD = 1.07$). The average relative proportion of interethnic dating per school, as a second relative measure of the descriptive norm, was equal to 2.23% ($SD = 2.78$). The respondents had an average age of 203.9 months ($SD = 7.77$), with 66% reporting being girls and 34% being boys. In terms of ethnic composition, 84% of respondents identified as ethnic majority members, while 16% identified as ethnic minority members.

Table 2 presents Pearson correlations for the key theoretical variables in the observed sample. In line with our theoretical expectations, interethnic dating was significantly positively correlated with individual attitudes ($r = .274$) and the average ingroup attitude ($r = .488$). However, contrary to expectations, it was not significantly associated with either absolute ($r = .023$) or relative descriptive norms ($r = .065$) and was significantly negatively correlated with the average outgroup attitude ($r = -.333$).

Table 2. Pearson correlations for key theoretical variables in the observed sample.

	Inter. date	Ind. att.	Ing. att.	Outg. att.	Desc. norm. (abs.)	Desc. norm. (rel.)	Eth. status
Inter. date		0.274***	0.488***	−0.333***	0.023	0.065	0.610***
Ind. att.			0.425***	−0.080	0.007	0.044	0.264***
Ing. att.				−0.270***	0.036	0.095*	0.650***
Outg. att.					0.050	−0.018	−0.478***
Desc. norm (abs.)						0.816***	0.110*
Desc. norm (rel.)							0.127**
Eth. status							

Notes: *** $p < .001$; ** $p < .01$; * $p < .05$, two-sided.

3.6. Analytical Strategy

To examine our hypotheses, we employed a two-step analytical strategy. First, we estimated a logistic regression model (i.e., the outcome model) to evaluate how the likelihood of being in an interethnic romantic relationship in either wave 2 or 3 was associated with our independent variables. We began with a model including only main effects and then introduced interaction terms to test our moderation hypotheses. Specifically, we examined whether the association between individual attitudes and interethnic dating was moderated by the average outgroup attitude and by descriptive norms by including interaction terms between each of these variables and individual attitudes. All continuous variables were mean-centered prior to analysis to facilitate interpretation of interaction terms. Throughout the analyses, we present separate models in which descriptive norms are operationalized in either absolute or relative terms. To probe significant interactions, we used Johnson-Neyman plots to identify the regions of significance along the moderator, allowing us to determine the values of the moderator at which individual attitudes significantly predicted interethnic dating. To account for multiple testing across these conditional effects, we applied a correction for multiple comparisons.

To test the hypothesized mediation effect of the average ingroup attitude on interethnic dating via individual attitudes, we followed the causal mediation framework proposed by Tingley et al. (2014). This involved first estimating a linear regression model (i.e., the mediator model) with individual attitudes as the dependent variable and the average ingroup attitude as the main predictor, while controlling for the average outgroup attitude, descriptive norm, and the control variables. The results of the mediator models are available in the Supplementary File. Because the outcome model included interaction terms, the resulting mediation effects represent conditional indirect effects, with the moderators (i.e., the average outgroup attitude and descriptive norm) fixed at their means. We then conducted a bootstrapped mediation analysis using the mediation package in R, based on 2,000 simulations, incorporating the mediator and outcome models.

Given the hierarchical structure of the data—with adolescents nested within schools—we adjusted for the intra-class correlation by estimating cluster-robust standard errors in all models. For analyses involving the imputed data, results were pooled across the 40 imputed datasets using Rubin's rules with the Barnard and Rubin (1999) adjustment for degrees of freedom. The degrees of freedom were however greatly reduced in

the imputed sample, even though the number of cases more than doubled. This lowers confidence in results obtained for the imputed sample. All analyses with respect to the imputed sample (mirroring those of the observed sample) are therefore only reported in the Supplementary File.

4. Results

Table 3 presents the results of the logistic regression models predicting interethnic dating based on the theoretical and control variables. For reasons of modelling order, we start with hypotheses 2 and 3. Our second hypothesis posited that the average outgroup attitude towards the respondent's ingroup would moderate the association between individual attitudes towards the ethnic outgroup and the likelihood of engaging in interethnic dating. Specifically, we expected this association to be stronger when the average outgroup attitude was more positive. However, the interaction between individual attitudes and average outgroup attitudes was not statistically significant in either model 1b (β = .00, p = .55) or 2b (β = .00, p = .55), providing no empirical support for this hypothesis.

Our third hypothesis proposed that descriptive norms encouraging interethnic relationships would moderate the association between individual attitudes towards the ethnic outgroup and the likelihood of engaging in interethnic dating, such that this association would be stronger when descriptive norms indicated that interethnic relationships were more common. First, regarding the marginal effect of individual attitudes, model 1b shows a statistically significant positive association when absolute descriptive norms (i.e., number of interethnic relationships)were at their mean (β = .02, p = .04). In contrast, in model 2b, this marginal effect was not statistically significant when descriptive norms were operationalized in relative terms (i.e., the number relative to the potential number of interethnic relationships; β = .02, p = .10). Subsequent inspection of the interaction terms between descriptive norms and individual attitudes in Table 3 supports the hypothesis for both operationalizations (1b: β = .02, p = .01; 2b: β = .01, p = .004). In both models, the positive interaction terms indicate that more positive individual attitudes were more predictive of interethnic dating in normative contexts where interethnic relationships were perceived as more common. Conversely, more negative attitudes were associated with a lower likelihood of interethnic dating in these same, more supportive contexts.

Turning to the other side of the interaction, model 1b shows a significant negative marginal effect of absolute descriptive norms (β = −.43, p = .02), indicating that when individual attitudes are average, stronger descriptive norms are associated with a lower likelihood of interethnic dating. In contrast, relative descriptive norms did not show a significant marginal effect in model 2b (β = −.10, p = .11). In both cases, however, these marginal effects are qualified by the significant positive interaction with individual attitudes: the more positive the individual attitude, the weaker the negative effect of descriptive norms.

To further clarify the nature of this interaction, we examined the Johnson-Neyman plots shown in Figures 2 and 3. These plots illustrate how the effect of individual attitudes on interethnic dating varies across the observed range of descriptive norms (Figure 2), and how the effect of descriptive norms varies across the range of individual attitudes (Figure 3). In Figure 2, for the absolute descriptive norm, we probed the marginal effect of individual attitudes at values 0, 1, 2, 3, and 4 on the original scale. For the relative norm, we selected the minimum and maximum observed values, along with three equally spaced values in between. In Figure 3, we probed the marginal effects of both absolute and relative descriptive norms across the minimum and

Table 3. Logistic regression model of interethnic dating on independent theoretical and control variables, by absolute (model 1) and relative (model 2) descriptive norm operationalization ($N = 511$).

	(1a)		(1b)		(2a)		(2b)	
	Est. (se)		Est. (se)		Est. (se)		Est. (se)	
Intercept	−1.63*	(.696)	−1.47*	(.702)	−1.68*	(.736)	−1.49*	(.724)
Individual attitudes	.02*	(.009)	.02*	(.010)	.02*	(.009)	.02	(.011)
Ingroup attitudes	−.01	(.031)	−.01	(.032)	−.01	(.032)	−.00	(.033)
Outgroup attitudes	.00	(.016)	.00	(.016)	−.00	(.017)	.00	(.017)
Descriptive norms (abs.)	−.24	(.171)	−.43*	(.188)				
Descriptive norms (rel.)					−.04	(.055)	−.10	(.064)
Traditional gender roles	−.02	(.158)	−.01	(.150)	−.03	(.154)	−.04	(.152)
Conservative family values	.04	(.277)	.01	(.321)	.02	(.288)	−.04	(.325)
Mixed parents	.60	(.598)	.54	(.664)	.63	(.599)	.50	(.626)
Educated parents	−1.36	(.753)	−1.72*	(.805)	−1.29	(.791)	−1.65*	(.800)
Religious affiliation								
No religion (ref)								
Christianity	−1.11*	(.535)	−1.03*	(.518)	−1.08*	(.522)	−1.01*	(.503)
Islam	−1.57	(.863)	−1.28	(.924)	−1.46	(.854)	−1.26	(.908)
Other religion	−.62	(1.22)	−.30	(1.22)	−.58	(1.19)	−.32	(1.19)
Proportion outgroup	.04	(.023)	.04	(.023)	.04	(.023)	.04	(.023)
Age (in months)	.05*	(.025)	.06*	(.026)	.05*	(.026)	.06*	(.027)
Gender (1 = girl)	.27	(.401)	.39	(.447)	.26	(.411)	.39	(.439)
Ethnic majority/minority status (1 = minority)	3.13***	(.966)	2.86**	(.974)	3.02**	(.951)	2.86**	(.960)
Outg. att.*Ind. att.			.00	(.001)			.00	(.001)
Desc. norm (abs.)*Ind. att.			.02*	(.009)				
Desc. norm (rel.)*Ind. att.							.01**	(.003)

Notes: *** $p < .001$; ** $p < .01$; * $p < .05$, two-sided.

maximum observed values of individual attitudes, and three equally spaced values in between. The bottom x-axes display centered values; the top x-axes show the original scales to aid interpretation.

Figure 2 reveals that for both the absolute and relative norm operationalizations, individual attitudes significantly predicted interethnic dating only when descriptive norms were sufficiently supportive. For the absolute norm, this threshold was crossed when there was at least one observed interethnic relationship in the school. For the relative norm, significance emerged when roughly 3% of peers were in interethnic relationships. Below these thresholds, individual attitudes were not significantly associated with interethnic dating. Figure 3 shows that absolute descriptive norms had a significantly negative marginal effect on

interethnic dating until individual attitudes were approximately at the midpoint of the attitude scale (i.e., neutral attitudes). Beyond that point, the effect was no longer statistically significant. For relative descriptive norms, no significant marginal effect was observed across the full range of individual attitudes.
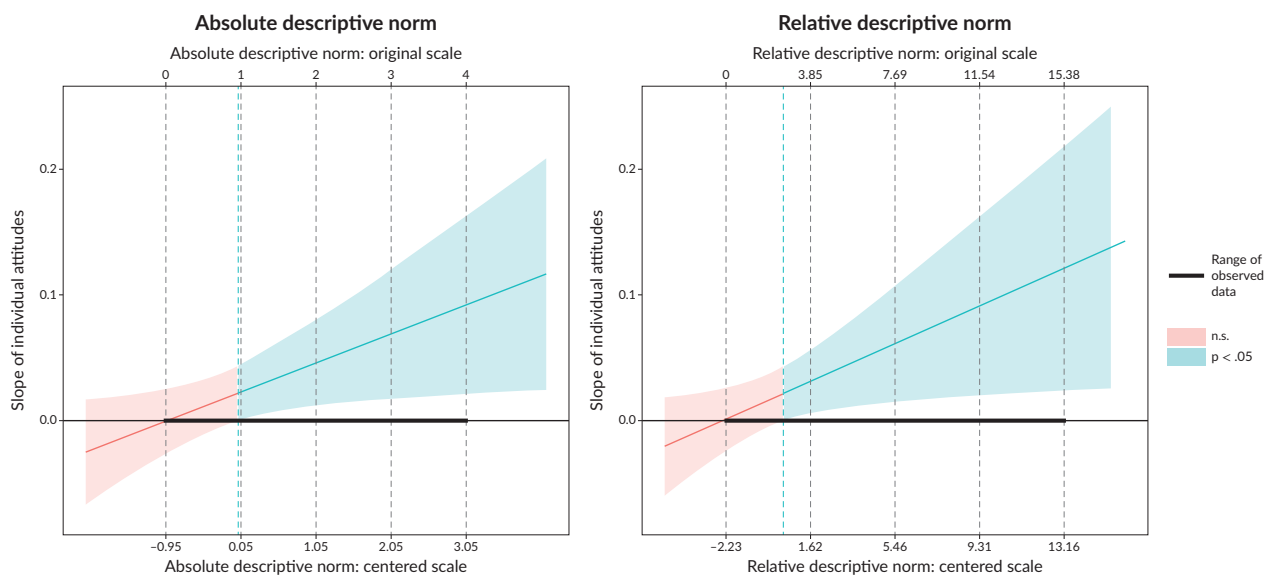
Figure 2. Johnson-Neyman plot of the marginal effect of individual attitudes on interethnic dating conditional on absolute and relative descriptive norms.
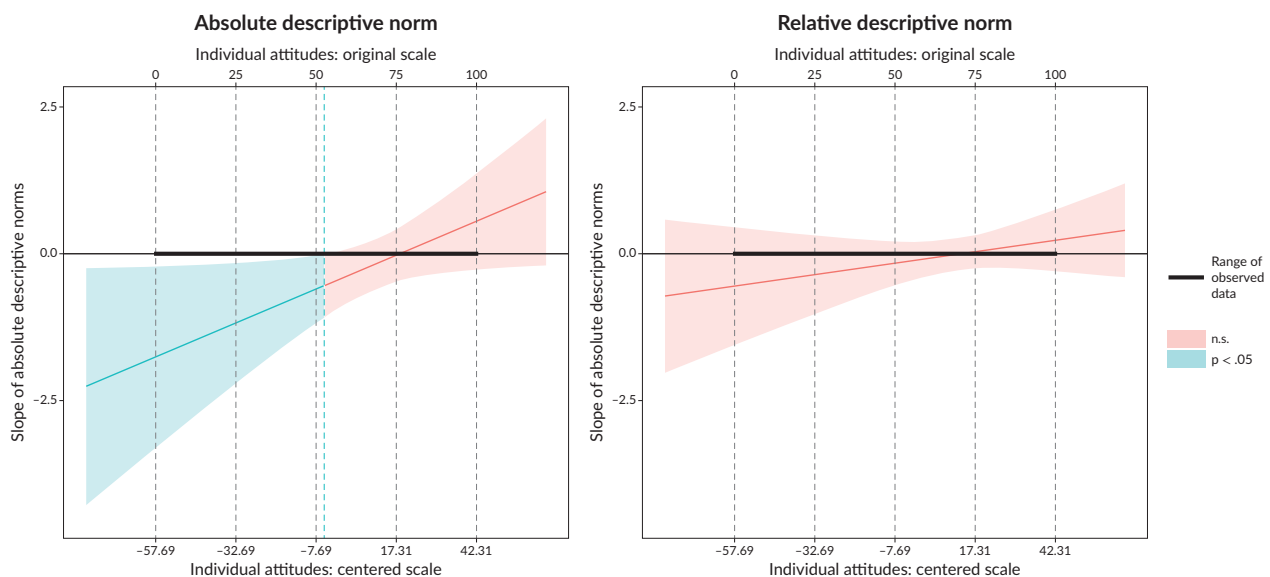
Figure 3. Johnson-Neyman plot of the marginal effect of absolute and relative descriptive norms on interethnic dating conditional on individual attitudes.

Taken together, this interaction pattern suggests that individuals with more positive attitudes towards the ethnic outgroup are increasingly able to translate these attitudes into interethnic dating as the normative environment becomes sufficiently and increasingly supportive. However, counterintuitively, individuals with more negative attitudes, are likewise better able to translate their attitudes into a lower likelihood of interethnic dating under similarly supportive normative conditions. In addition, for individuals with

(sufficiently) negative attitudes, unsupportive absolute descriptive norms show a negative marginal effect on interethnic dating. Notably, we do not observe a statistically significant negative effect of absolute descriptive norms when individual attitudes are positive, even in unsupportive normative environments.

With respect to the control variables, Table 3 shows that having educated parents emerges as a negative predictor in the interaction models: respondents with highly educated parents are less likely to date interethnically (1b: $\beta = -1.72$, $p = .03$; 2b: $\beta = -1.65$, $p = .04$). Religious affiliation also appears to play an important role in shaping interethnic dating patterns, with individuals identifying as Christian being significantly less likely to engage in interethnic dating compared to those with no religious affiliation (1b: $\beta = -1.03$, $p = .047$; 2b: $\beta = -1.01$, $p = .045$). Age is furthermore positively associated with interethnic dating: older respondents are more likely to report being in an interethnic relationship (1b: $\beta = .06$, $p = .03$; 2b: $\beta = .06$, $p = .04$). Finally, ethnic group membership stands out as a strong predictor of interethnic dating: ethnic minority individuals are significantly more likely than ethnic majority individuals to report interethnic dating (1b: $\beta = 2.86$, $p = .003$; 2b: $\beta = 2.86$, $p = .003$).

Table 4 finally presents the results of the mediation model testing our first hypothesis. We expected that more positive average ingroup attitudes towards ethnic outgroups would increase positive individual attitudes towards ethnic outgroups, which in turn would be associated with a greater likelihood of engaging in interethnic dating. Inspection of Table 4 indicates that this expectation was supported when descriptive norms were operationalized in absolute terms ((1) $\beta = .0012$, $p = .04$), but not when operationalized in relative terms ((2) $\beta = .0010$, $p = .10$). Additionally, we do not observe significant direct ((1): $\beta = -.0002$, $p = .87$; (2): $\beta = -.0002$, $p = .90$) or total effects ((1): $\beta = -.0010$, $p = .66$; (2): $\beta = .0009$, $p = .74$) in either model.

Table 4. Mediation model for interethnic dating on average ingroup attitudes via individual attitudes, by absolute and relative descriptive norm operationalization.

	(1) Descriptive norm: Absolute	(2) Descriptive norm: Relative
	Est. [95% C.I.]	Est. [95% C.I.]
Indirect effect	**.0012* [.0001; .0025]**	.0010 [−.0002; .0024]
Direct effect	−.0002 [−.0039; .0046]	−.0002 [−.0039; .0045]
Total effect	.0010 [−.0027; .0056]	.0009 [−.0027; .0055]

Notes: *** $p < .001$; ** $p < .01$; * $p < .05$, two-sided. Indirect effects and confidence intervals reflect conditional indirect effects, with moderators fixed at their means.

5. Conclusions and Discussion

In this study, we investigated the interethnic partner choice of adolescents. Using unique large-scale panel data collected among adolescents with and without a migration background in the Netherlands, we tested hypotheses about the influence of in- and outgroup attitudes and descriptive norms on interethnic dating patterns.

First, in support of our first hypothesis, we found that the average ingroup attitude had a positive indirect effect via personal preferences for interethnic dating on the likelihood of interethnic dating in the model where descriptive norms were operationalized as absolute. Second, individual attitudes towards the ethnic

outgroup predicted interethnic dating at a later point in time and this association was moderated by both the absolute and relative operationalizations of the descriptive norm, but not by the average attitude of the ethnic outgroup towards the ingroup. The second hypothesis, which stated that the relationship between personal preference for interethnic dating and the likelihood of engaging in such dating would be stronger when outgroup attitudes towards the ingroup were more positive, was therefore not confirmed. We did find support for our third hypothesis: descriptive norms strengthen the link between personal preferences and interethnic dating, with this relationship becoming stronger as interethnic dating becomes more common in the school environment. With respect to the control variables, we found that both ethnic minority adolescents and older students were more likely to date someone from an ethnic outgroup, while students with at least one biological or adoptive parent who had completed upper secondary school, and Christian students compared to non-religious students, were less likely to do so.

A particularly interesting finding of this study was the interaction effect between individual attitudes towards the ethnic outgroup and descriptive norms on the likelihood of interethnic dating. We found that individuals with more positive attitudes towards the ethnic outgroup were increasingly able to translate these attitudes into interethnic dating as the descriptive normative environment became sufficiently and increasingly supportive. However, counterintuitively, individuals with more negative attitudes were likewise better able to translate their attitudes into a lower likelihood of interethnic dating under such supportive normative conditions. In addition, for individuals with sufficiently negative attitudes, unsupportive absolute but not relative descriptive norms showed a negative marginal effect on interethnic dating. Notably, we did not observe a statistically significant negative effect of absolute descriptive norms when individual attitudes were (sufficiently) positive, even in unsupportive normative environments.

As such, our findings show that within this particular context, supportive normative environments do not universally promote interethnic dating but instead amplify the effect of individuals' existing attitudes. This qualifies the role of descriptive norms in shaping interethnic dating: rather than exerting a uniformly liberalizing influence, they appear to facilitate the expression of both openness and resistance to interethnic relationships, depending on individual predispositions. Additionally, we found that more positive average ingroup attitudes were indirectly associated with a higher likelihood of interethnic dating via stronger personal preferences, suggesting that the attitudinal climate within an ingroup may contribute to shaping individual-level openness to interethnic dating. These results underscore the importance of considering not only individual attitudes in shaping interethnic partner choice, but also their interplay with the local normative environment.

The lack of support for the influence of the attitudes of outgroup peers within the highly-relevant school context was surprising. A potential explanation for the lack of effects of outgroup peer attitudes towards the (relevant) ingroups is that we measured actual attitudes rather than perceived attitudes. Adolescents may perceive the attitudes of their outgroup peers incorrectly, and be guided by these misperceptions. For example, adolescents may erroneously conclude that interethnic relationships are disapproved of by their outgroup peers, thereby lowering their likelihood of interethnic dating. Another explanation for the lack of support might be that we did not directly assess outgroup peers' preferences for interethnic dating, but rather measured outgroup peers' general attitudes towards the ingroup. While this measure was a pragmatic consequence of the data available—and research shows that such outgroup peer attitudes are in fact associated with popularity and likability (van Vemde et al., 2023) and with visible behaviors such as intergroup contact (Pettigrew & Tropp, 2008; Swart et al., 2011)—by using a general attitude measure, we

may have overestimated outgroup peers' openness and approval of interethnic dating. People may differentiate between their preferences for outgroup relationships in the public or the private sphere and are generally more positive regarding relations in the public sphere (Bogardus, 1925, 1933; Ng Tseung-Wong & Verkuyten, 2015). If outgroup peers' general attitude towards an ethnic group is negative, their approval of interethnic romantic relationships with members of that group is almost certainly negative as well. But when general attitudes towards an ethnic group become more positive, the possibility of approval for interethnic dating increases, and more so if attitudes become very positive. This is also demonstrated by the finding that adolescents' own attitudes towards the outgroup significantly predicted their likelihood of engaging in an interethnic relationship, prior to the inclusion of interaction terms. By using the general attitude of outgroup peers as a measure, we argue that we may have thus estimated the lower bound of the true effect.

A strength of this study is that interethnic partner choice was studied from a two-sided perspective, both theoretically and empirically. Attitudes regarding the outgroup were measured within a relevant context (i.e., the school as a local dating market) and reflected the majority and minority group. Descriptive norms were operationalized in a manner similar to how adolescents would perceive them. Nevertheless, this study comes with limitations.

Unfortunately, we were not able to test for differences in the effects between ethnic minority origin groups because of the small group sizes. Although differences between these groups are partly captured by the differences in natives' attitudes towards one's origin group and other covariates, some differences likely remained. Longitudinal data with a larger sample (including other ethnic minority groups) is needed to gain insight into how group differences may have affected the findings.

A further limitation concerns our operationalization of descriptive norms as either absolute or relative. When using the absolute version, we found an indirect effect of average ingroup attitudes on interethnic dating via individual attitudes; this effect was not observed with the relative version. The absolute and relative norms also showed slightly different moderating effects on the relationship between individual attitudes and interethnic dating. This may reflect a substantive distinction between the two: absolute norms capture the total number of interethnic relationships within a school, whereas relative norms indicate how common interethnic dating is among all romantic relationships, thereby adjusting for school size and relationship prevalence. Alternatively, these differences may be artefacts of how the measures were constructed. Future research should examine how adolescents interpret and weigh these different normative cues when shaping their interethnic dating behaviours.

A third limitation of this study is that we were unable to control for non-random selection into schools. Immigrants who are very negative about natives probably select themselves into schools with a low proportion of natives, and natives who are very negative about immigrants probably select themselves into schools with a low proportion of immigrants. The observed null effect of the proportion of outgroup students at school may therefore be biased upwards.

Other sources of normative cues on interethnic dating, relevant to adolescents, were also beyond the scope of this article, such as (social) media or associations they belong to, but should be considered in future research. Finally, we cannot say how stable these interethnic relations were, and whether they will increase the likelihood of ethnic intermarriage in the future for the adolescents we studied. How interethnic dating in adolescence impacts overall integration in society thus remains a somewhat open question.

In sum, we conclude that personal preferences, ingroup attitudes, and descriptive peer norms jointly shape interethnic dating patterns in Dutch high schools in the studied period, next to demographic and parental influences. At the age of the adolescents we study, we find no evidence that the (perhaps privately held or misperceived) attitudes of their outgroup peers within school about interethnic relations have an influence. Altogether, our findings strengthen the theoretical model on preferences, opportunities, and third-party influences by showing how the local context shapes whether individual preferences are translated into romantic interethnic dating behaviours.

Conflict of Interests
The authors declare no conflict of interests. In this article, editorial decisions were undertaken by Ulf R. Hedetoft (University of Copenhagen, Denmark).

Data Availability
The R-code that supports the findings of this study is available at https://osf.io/2zya9. The data are available upon request to other scientists. Inquiries will be forwarded to the relevant data depositories.

Supplementary Material
Supplementary material for this article is available online in the format provided by the author (unedited).

References
Alba, R., & Nee, V. (2003). *Remaking the American mainstream: Assimilation and contemporary integration*. Harvard University Press.

Azjen, I., Fishbein, M., Lohman, S., & Albarracín, D. (2018). The influence of attitudes on behavior. In D. Albarracín & B. T. Johnson (Eds.), *The handbook of attitudes* (2nd ed., pp. 197–255). Routledge.

Barnard, J., & Rubin, D. B. (1999). Miscellanea. Small-sample degrees of freedom with multiple imputation. *Biometrika*, *86*(4), 948–955. https://doi.org/10.1093/biomet/86.4.948

Bayram Özdemir, S., Sun, S., Korol, L., Özdemir, M., & Stattin, H. (2018). Adolescents' engagement in ethnic harassment: Prejudiced beliefs in social networks and classroom ethnic diversity. *Journal of Youth and Adolescence*, *47*(6), 1151–1163. https://doi.org/10.1007/s10964-017-0795-0

Blackwell, D. L., & Lichter, D. T. (2004). Homogamy among dating, cohabiting and married couples. *The Sociological Quarterly*, *45*(4), 719–737. https://doi.org/10.1111/j.1533-8525.2004.tb02311.x

Blau, P. M. (1977). A macrosociological theory of social structure. *American Journal of Sociology*, *83*(1), 26–54. https://doi.org/10.1086/226505

Bogardus, E. S. (1925). Social distance and its origins. *Journal of Applied Sociology*, *9*, 216–226.

Bogardus, E. S. (1933). A social distance scale. *Sociology and Social Research*, *17*, 265–271.

Bourgeois, K. S., & Leary, M. R. (2001). Coping with rejection: Derogating those who choose us last. *Motivation and Emotion*, *25*(2), 101–111. https://doi.org/10.1023/a:1010661825137

Cialdini, R. B., Kallgren, C. A., & Reno, R. R. (1991). A focus theory of normative conduct: A theoretical refinement and reevaluation of the role of norms in human behavior. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 24, pp. 201–234). Academic Press. https://doi.org/10.1016/S0065-2601(08)60330-5

Cillessen, A. H. N., & Rose, A. J. (2005). Understanding popularity in the peer system. *Current Directions in Psychological Science*, *14*(2), 102–105. https://doi.org/10.1111/j.0963-7214.2005.00343.x

Coleman, D. (2006). Immigration and ethnic change in low-fertility countries: A third demographic transition. *Population and Development Review*, *32*(3), 401–446. https://doi.org/10.1111/j.1728-4457.2006.00131.x

Connolly, J., Craig, W., Goldberg, A., & Pepler, D. (2004). Mixed-gender groups, dating, and romantic relationships in early adolescence. *Journal of Research on Adolescence*, *14*(2), 185–207. https://doi.org/10.1037//0033–2909.111.1.62

Connolly, J., Nguyen, H. N. T., Pepler, D., Craig, W., & Jiang, D. (2013). Developmental trajectories of romantic stages and associations with problem behaviours during adolescence. *Journal of Adolescence*, *36*(6), 1013–1024. https://doi.org/10.1016/j.adolescence.2013.08.006

de Vroome, T., Verkuyten, M., & Martinovic, B. (2014). Host national identification of immigrants in the Netherlands. *International Migration Review*, *48*(1), 76–102. https://doi.org/10.1111/imre.12063

Dollmann, J., Jacob, K., & Kalter, F. (2014). *Examining the diversity of youth in Europe: A classification of generations and ethnic origins using CILS4EU data (technical report)* (Working Paper). MZES. https://access.gesis.org/dbk/56516

Durkheim, E. (1951). *Suicide: A study in sociology* (2nd ed.). Routledge.

Feliciano, C., & Hijara, C. M. (2024). Different but somewhat similar: Panethnicity, group boundaries, and dating preferences among Asian and Latino college students. *Sociology of Race and Ethnicity*, *10*(3), 353–370. https://doi.org/10.1177/23326492231222175

Finnäs, F., & O'Leary, R. (2003). Choosing for the children: The affiliation of the children of minority-majority group intermarriages. *European Sociological Review*, *19*(5), 483–499. https://doi.org/10.1093/esr/19.5.483

Fujino, D. C. (1997). The rates, patterns and reasons for forming heterosexual interracial dating relationships among Asian Americans. *Journal of Social and Personal Relationships*, *14*(6), 809–828. https://doi.org/10.1177/0265407597146006

Grusec, J. E., & Hastings, P. D. (Eds.). (2014). *Handbook of socialization: Theory and research*. Guilford Publications.

Harris, D. R., & Ono, H. (2005). How many interracial marriages would there be if all groups were of equal size in all places? A new look at national estimates of interracial marriage. *Social Science Research*, *34*(1), 236–251. https://doi.org/10.1016/j.ssresearch.2004.01.002

Heckman, J. J. (1979). Sample selection bias as a specification error. *Econometrica*, *47*(1), 153–161. https://doi.org/10.2307/1912352

Herman, M. R., & Campbell, M. E. (2012). I wouldn't, but you can: Attitudes toward interracial relationships. *Social Science Research*, *41*(2), 343–358. https://doi.org/10.1016/j.ssresearch.2011.11.007

Huijnk, W., & Liefbroer, A. C. (2012). Family influences on intermarriage attitudes: A sibling analysis in the Netherlands. *Journal of Marriage and Family*, *74*(1), 70–85. https://doi.org/10.1111/j.1741-3737.2011.00882.x

Huijnk, W., Verkuyten, M., & Coenders, M. (2012). Family relations and the attitude towards ethnic minorities as close kin by marriage. *Ethnic and Racial Studies*, *36*(11), 1890–1909. https://doi.org/10.1080/01419870.2012.699086

Kalmijn, M. (1998). Intermarriage and homogamy: Causes, patterns, trends. *Annual Review of Sociology*, *24*, 395–421. https://doi.org/10.1146/annurev.soc.24.1.395

Kalmijn, M. (2015). The children of intermarriage in four European countries: Implications for school achievement, social contacts, and cultural values. *The ANNALS of the American Academy of Political and Social Science*, *662*(1), 246–265. https://doi.org/10.1177/0002716215595391

Kalmijn, M., & Flap, H. (2001). Assortative meeting and mating: Unintended consequences of organized settings for partner choices. *Social Forces*, *79*(4), 1289–1312. http://www.jstor.org/stable/pdfplus/2675473.pdf

Kalmijn, M., & van Tubergen, F. (2010). A comparative perspective on intermarriage: Explaining differences among national-origin groups in the United States. *Demography*, *47*(2), 459–479. https://doi.org/10.1353/dem.0.0103

Kalter, F., Heath, A. F., Hewstone, M., Jonsson, J. O., Kalmijn, M., Kogan, I., & van Tubergen, F. (2014). *Children of immigrants longitudinal survey in four European countries (CILS4EU)—Full version* (GESIS Data Archive ZA5353; Version 1.1.0) [Data set]. GESIS Data Archive. https://doi.org/10.4232/cils4eu.5353.1.1.0

Kalter, F., Heath, A. F., Hewstone, M., Jonsson, J. O., Kalmijn, M., Kogan, I., & van Tubergen, F. (2016a). *Children of immigrants longitudinal survey in four European countries (CILS4EU)—Full version* (GESIS Data Archive ZA5353; Version 1.0.0) [Data set]. GESIS Data Archive. https://doi.org/10.4232/cils4eu.5353.1.0.0

Kalter, F., Heath, A. F., Hewstone, M., Jonsson, J. O., Kalmijn, M., Kogan, I., & van Tubergen, F. (2016b). *Children of immigrants longitudinal survey in four European countries (CILS4EU)—Full version* (GESIS Data Archive ZA5353; Version 1.2.0) [Data set]. GESIS Data Archive. https://doi.org/10.4232/cils4eu.5353.1.2.0

Kalter, F., Heath, A. F., Hewstone, M., Jonsson, J. O., Kalmijn, M., Kogan, I., & van Tubergen, F. (2016c). *Children of immigrants longitudinal survey in four European countries (CILS4EU)—Full version* (GESIS Data Archive ZA5353; Version 2.3.0) [Data set]. GESIS Data Archive. https://doi.org/10.4232/cils4eu.5353.2.3.0

Kalter, F., Heath, A. F., Hewstone, M., Jonsson, J. O., Kalmijn, M., Kogan, I., & van Tubergen, F. (2016d). *Children of immigrants longitudinal survey in four European countries (CILS4EU)—Full version* (GESIS Data Archive ZA5353; Version 3.1.0) [Data set]. GESIS Data Archive. https://doi.org/10.4232/cils4eu.5353.3.1.0

Kiesner, J., Kerr, M., & Stattin, H. (2004). "Very important persons" in adolescence: Going beyond in-school, single friendships in the study of peer homophily. *Journal of Adolescence*, *27*(5), 545–560. https://doi.org/10.1016/j.adolescence.2004.06.007

King, R. B., & Harris, K. M. (2007). Romantic relationships among immigrant adolescents. *International Migration Review*, *41*(2), 344–370. https://doi.org/10.1111/j.1747-7379.2007.00071.x

König, S., Stang-Rabrig, J., & McElvany, N. (2022). Adolescents' implicit attitudes towards people with immigrant background: Differences and correlates. *Social Psychology of Education*, *25*(6), 1381–1409. https://doi.org/10.1007/s11218-022-09722-6

Lewis, K. (2016). Preferences in the early stages of mate choice. *Social Forces*, *95*(1), 283–320. https://doi.org/10.1093/sf/sow036

Lundquist, J., Lin, K. H., & Curington, C. (2024). Interracial unions and racial assortative mating in an age of growing diversity, shifting intimate relationships, and emerging technologies. *Annual Review of Sociology*, *50*, 431–453. https://doi.org/10.1146/annurev-soc-083023-023534

Madsen, S. D., & Collins, W. A. (2011). The salience of adolescent romantic experiences for romantic relationship qualities in young adulthood. *Journal of Research on Adolescence*, *21*(4), 789–801. https://doi.org/10.1111/j.1532-7795.2011.00737.x

Martinović, B. (2013). The inter-ethnic contacts of immigrants and natives in the Netherlands: A two-sided perspective. *Journal of Ethnic and Migration Studies*, *39*(1), 69–85. http://doi.org/10.1080/1369183X.2013.723249

McPherson, M., Smith-Lovin, L., & Cook, J. M. (2001). Birds of a feather: Homophily in social networks. *Annual Review of Sociology*, *27*, 415–444. https://www.jstor.org/stable/2678628

Merton, R. K. (1941). Intermarriage and the social structure. *Psychiatry*, *4*(3), 361–374. https://doi.org/10.1080/00332747.1941.11022354

Mollenhorst, G., Völker, B., & Flap, H. (2008). Social contexts and personal relationships: The effect of meeting opportunities on similarity for relationships of different strength. *Social Networks*, *30*(1), 60–68. https://doi.org/10.1016/j.socnet.2007.07.003

Moody, J. (2001). Race, school integration, and friendship segregation in America. *American Journal of Sociology*, *107*(3), 679–716. https://doi.org/10.1086/338954

Munniksma, A., Flache, A., Verkuyten, M., & Veenstra, R. (2012). Parental acceptance of children's intimate ethnic outgroup relations: The role of culture, status, and family reputation. *International Journal of Intercultural Relations*, *36*(4), 575–585. https://doi.org/10.1016/j.ijintrel.2011.12.012

Ng Tseung-Wong, C., & Verkuyten, M. (2015). "I'd rather we be neighbours than lovers": The two-sidedness of multiculturalism. *Group Processes & Intergroup Relations*, *18*(4), 437–453. https://doi.org/10.1177/1368430214546068

Pehar, L., Čorkalo Biruški, D., & Pavin Ivanec, T. (2020). The role of peer, parental, and school norms in predicting adolescents' attitudes and behaviours of majority and different minority ethnic groups in Croatia. *PLOS ONE*, *15*(1), Article 0227512. https://doi.org/10.1371/journal.pone.0227512

Pettigrew, T. F. (1998). Intergroup contact theory. *Annual Review of Psychology, 49*, 65–85. http://arjournals.annualreviews.org/doi/pdf/10.1146/annurev.psych.49.1.65

Pettigrew, T. F., & Tropp, L. R. (2008). How does intergroup contact reduce prejudice? Meta-analytic tests of three mediators. *European Journal of Social Psychology*, *38*(6), 922–934. https://doi.org/10.1002/ejsp.504

Qian, Z. (2004). Options: Racial/ethnic identification of children of intermarried couples. *Social Science Quarterly*, *85*(3), 746–766. https://doi.org/10.1111/j.0038-4941.2004.00243.x

Raley, R. K., Crissey, S., & Muller, C. (2007). Of sex and romance: Late adolescent relationships and young adult union formation. *Journal of Marriage and Family*, *69*(5), 1210–1226. https://doi.org/10.1111/j.1741-3737.2007.00442.x

Rivas-Drake, D., Umaña-Taylor, A. J., Schaefer, D. R., & Medina, M. (2017). Ethnic-racial identity and friendships in early adolescence. *Child Development*, *88*(3), 710–724. https://doi.org/10.1111/cdev.12790

Shibazaki, K., & Brennan, K. A. (1998). When birds of different feathers flock together: A preliminary comparison of intra-ethnic and inter-ethnic dating relationships. *Journal of Social and Personal Relationships*, *15*(2), 248–256. https://doi.org/10.1177/0265407598152007

Swart, H., Hewstone, M., Christ, O., & Voci, A. (2011). Affective mediators of intergroup contact: A three-wave longitudinal study in South Africa. *Journal of Personality and Social Psychology*, *101*(6), 1221–1238. https://doi.org/10.1037/a0024450

Thijs, J., Verkuyten, M., & Grundel, M. (2014). Ethnic classroom composition and peer victimization: The moderating role of classroom attitudes. *Journal of Social Issues*, *70*(1), 134–150. https://doi.org/10.1111/josi.12051

Tingley, D., Yamamoto, T. Hirose, K., Keele, L., & Imai, K. (2014). Mediation: R Package for causal mediation analysis. *Journal of Statistical Software*, *59*(5), 1–38. http://www.jstatsoft.org/v59/i05

Tropp, L. R., & Pettigrew, T. F. (2005). Relationships between intergroup contact and prejudice among minority and majority status groups. *Psychological Science*, *16*(12), 951–957. https://doi.org/10.1111/j.1467-9280.2005.01643

van Buuren, S. (2018). *Flexible imputation of missing data* (2nd ed.). Chapman and Hall/CRC. https://doi.org/10.1201/9780429492259

van Buuren, S., & Groothuis-Oudshoorn, K. (2011). Mice: Multivariate imputation by chained equations in R. *Journal of Statistical Software*, *45*, 1–67. https://doi.org/10.18637/jss.v045.i03

van Tubergen, F., & Maas, I. (2007). Ethnic intermarriage among immigrants in the Netherlands: An analysis of population data. *Social Science Research*, *36*(3), 1065–1086. https://doi.org/10.1016/j.ssresearch.2006.09.003

van Vemde, L., Hornstra, L., & Thijs, J. (2021). Classroom predictors of national belonging: The role of interethnic contact and teachers' and classmates' diversity norms. *Journal of Youth and Adolescence*, *50*(8), 1709–1725. https://10.1007/s10964-021-01430-2

van Vemde, L., Thijs, J., & Hornstra, L. (2023). The classroom social environment in mixed secondary school classes: The role of ethnic classroom composition and classmates' explicitly and implicitly measured ethnic attitudes. *Social Psychology of Education*, *26*(2). https://doi.org/10.1007/s11218-022-09747-xvan

Zantvliet, P. I., Kalmijn, M., & Verbakel, E. (2014a). Early partner choices of immigrants: The effect of preferences, opportunities and parents on dating a native. *Journal of Ethnic and Migration Studies*, *41*(5), 1–23. https://doi.org/10.1080/1369183x.2014.948391

van Zantvliet, P. I., Kalmijn, M., & Verbakel, E. (2014b). Parental involvement in partner choice: The case of Turks and Moroccans in the Netherlands. *European Sociological Review*, *30*(3), 387–398. https://doi.org/10.1093/esr/jcu045

Vaquera, E., & Kao, G. (2008). Do you like me as much as I like you? Friendship reciprocity and its effects on school outcomes among adolescents. *Social Science Research*, *37*(1), 55–72. https://doi.org/10.1016/j.ssresearch.2006.11.002

Verbrugge, L. M. (1977). The structure of adult friendship choices. *Social Forces*, *56*(2), 576–597. https://doi.org/10.1093/sf/56.2.576

Wang, H., & Kao, G. (2007). Does higher socioeconomic status increase contact between minorities and Whites? An examination of interracial romantic relationships among adolescents. *Social Science Quarterly*, *88*(1), 146–164. https://doi.org/10.1111/j.1540-6237.2007.00451.x

Weißmann, M., & Maddox, A. (2016). *Early ethnic partner choice among natives and descendents of immigrants in Germany: The role of personal preferences and parental influence* (Working Paper). MZES. http://www.mzes.uni-mannheim.de/publications/wp/wp-164.pdf

Xie, Y., & Goyette, K. (1997). The racial identification of biracial children with one Asian parent: Evidence from the 1990 census. *Social Forces*, *76*(2), 547–570. https://doi.org/10.1093/sf/76.2.547

About the Authors

Eva Jaspers is a professor of empirical and theoretical sociology at Utrecht University. Her research is on gender, sexual, and racial/ethnic diversity from an interactional or network perspective. She aims to understand how people's interactions and relations with (dis)similar others shape their patterns of behavior and their outcomes at school, work, and at home.

Pascale van Zantvliet is a post-doctoral researcher affiliated with the Youth Education & Life Skills Community of Dynamics of Youth, Utrecht University. She holds a PhD in sociology and currently works at the Department of Education and Pedagogy. Her research interests include intergroup contact, perspective-taking, civic dialogue, citizenship education, and collaborative learning.

Jan-Willem Simons is a PhD student in the Department of Sociology at Utrecht University. He holds an MSc in applied statistics and a BSc in sociology, both from Utrecht University. His research focuses on interethnic relations, social networks, and applied statistical methods.

cogitatio

ARTICLE

Youth Norm Deviation and Intolerance: Pathways to Polarized Political Attitudes and Behavioral Intentions

Sebastian Lutterbach [ORCID] **and Andreas Beelmann** [ORCID]

Department of Research Synthesis, Intervention and Evaluation, Friedrich Schiller University Jena, Germany

Correspondence: Sebastian Lutterbach (sebastian.lutterbach@uni-jena.de)

Abstract

This study examines the psychological foundations of political polarization among adolescents, emphasizing the role of norm deviation and intolerance in shaping polarized political attitudes and behavioral intentions. We applied a structural equation model with latent variables to analyze the relationships between norm deviation and intolerance and five indicators of political polarization in 1,211 adolescents (aged 15–22 years, $M = 16.6$, $SD = 1.5$). These five indicators of political polarization were: rejection of democracy, rejection of human rights, conspiracy mentality, hostility toward foreigners, and the affinity for political violence. Multi-group structural equation modeling was conducted to explore the moderating effects of gender, age, and migration background. Results indicate that both norm deviation and intolerance are significantly associated with polarized political attitudes and behavioral intentions, with distinct patterns across demographic groups. These findings highlight the importance of tailored educational and preventive interventions that consider adolescents' gender, age, and migration background to reduce polarization risks and promote social cohesion.

Keywords

adolescence; intolerance; norm deviation; polarization; political attitudes

1. Introduction

Adolescence represents a critical developmental period for the formation of political attitudes and engagement (Neundorf & Smets, 2017; Sears & Brown, 2013). While a successful political socialization can promote prosocial convictions and democratic participation, this developmental phase also carries the risk of

extreme opinion formation, which may lead to societal divisions, conflicts, or alignment with radical political ideologies (Beelmann, 2020; Beelmann & Lutterbach, 2023; Berg-Schlosser et al., 2020). A multitude of psychological factors and social contexts shape the formation and polarization of political identities and beliefs during this life stage (Dekker et al., 2020; Wray-Lake, 2019). Norm deviation and intolerance constitute key factors in the development of political attitudes during adolescence. As young people seek to establish their own values and distance themselves from parental and institutional authorities, they often challenge social norms and use deviations as a means of exploring alternative lifestyles or ideologies (Eckstein et al., 2012; Quintelier, 2015). Simultaneously, the desire for peer recognition introduces a dynamic in which ingroup norms can either reinforce deviant behaviors or exacerbate intolerance toward outgroups perceived as norm-violating (Forst, 2013; Verkuyten et al., 2023).

To investigate how norm deviation and intolerance contribute to political polarization among adolescents, we examined their respective associations with key manifestations of polarized attitudes and behavioral intentions, including rejection of democracy and human rights, conspiracy mentality, hostility toward foreigners, and affinity for political violence. Additionally, we analyzed how gender, age, and migration background shape these relationships. This differential approach seeks to provide evidence-based insights into developmental and targeted strategies for preventing political polarization among youth (Beelmann, 2021; Beelmann & Lutterbach, 2022).

1.1. Political Polarization and Indicators of Polarized Political Attitudes and Behavioral Intentions

Polarization refers to the growing divide between individuals or social groups that fosters extreme viewpoints and reinforces ingroup and outgroup dynamics (Bakker & Lelkes, 2024). From a psychological perspective, polarization is defined as a process in which individuals' attitudes and beliefs become increasingly divergent, often accompanied by hostility toward opposing views (Lelkes, 2016). This phenomenon manifests in various forms, including the rejection of democratic principles, hostility toward political and social diversity, and the adoption of extremist ideologies. The consequences of polarization are profound, contributing to social fragmentation, the erosion of democratic institutions, and an increased risk of political violence (Bliuc et al., 2024; Piazza, 2023).

Polarization is particularly challenging during adolescence and early adulthood, a critical period for identity formation, norm and value development, and political socialization (Sears & Brown, 2013). At this stage, individuals are especially susceptible to social influences and may adopt polarized beliefs for the purpose of establishing and affirming their social identity. Research highlights that the interplay between developmental vulnerabilities and societal polarization can exacerbate phenomena such as ingroup superiority and political violence among youth (Rekker et al., 2015; Tyler & Iyengar, 2023). As adolescents navigate their social and political identities in polarized environments, they may gravitate toward extreme positions, further entrenching societal divides. This not only poses risks to their personal development but also threatens broader social cohesion.

From a social-psychological perspective, polarization manifests in attitudes such as rejection of democracy and human rights, conspiracy mentality, hostility toward foreigners, and affinity for political violence—each reflecting rigid group-based thinking. Social identity theory (Tajfel & Turner, 1979) suggests that identification with like-minded groups can heighten outgroup hostility and reduce openness to compromise.

For instance, conspiracy beliefs reinforce distrust and societal division by framing institutions or outgroups as hostile (Imhoff & Bruder, 2014). According to realistic group conflict theory (Sherif, 1966), perceived resource competition can fuel exclusionary attitudes, while threats to identity or status may foster authoritarian preferences and the erosion of democratic values (Osborne et al., 2023; Renström et al., 2022). In adolescence and early adulthood, such polarized attitudes are particularly salient due to identity formation and increased social influence (Crocetti et al., 2023; Wray-Lake, 2019). Political socialization theory and the impressionable years hypothesis (Sears & Levy, 2003) emphasize that early exposure to peer, family, and media cues, along with societal challenges, shape long-lasting political views. As youth adopt polarized attitudes—such as hostility toward foreigners or support for political violence—these can crystallize over time, embedding division into future political landscapes (Rekker et al., 2015).

1.2. Norm Deviation and Intolerance as Drivers of Political Polarization

1.2.1. Norm Deviation and Political Polarization

Norm deviation, understood as a departure from culturally and socially accepted standards and values (Hewstone et al., 2021) and thus rule-breaking behavior, reflects an early and general form or expression of polarization, as individuals distance themselves from widely shared societal norms. This deviation does not merely indicate a misalignment but can also function as a psychological and social mechanism that fosters openness to alternative worldviews and oppositional identities. In seeking new frameworks of meaning and belonging, norm-deviant individuals may adopt or reinforce politically polarized attitudes, which in turn contribute to broader processes of societal polarization.

The present study investigates norm deviation as a potential risk factor for political polarization. A central theoretical foundation is reactance theory (Brehm, 1966), which posits that individuals experience motivational reactance when they perceive their autonomy as unjustly constrained. Adolescents high in norm deviation may interpret dominant liberal-democratic values not as a shared framework but as illegitimate intrusions into their personal or group autonomy. This defiance against perceived restrictions can manifest as the rejection of democratic principles and human rights, especially when these are seen as limiting their behavioral or ideological space (Miron & Brehm, 2006). Further explanatory potential stems from social identity theory (Tajfel & Turner, 1979), which emphasizes the role of group membership in shaping attitudes and behaviors. Norm-deviating adolescents may feel excluded from or alienated by dominant normative peer groups and therefore seek belonging in alternative or oppositional ingroups that validate anti-system attitudes. This realignment of group identification can support the development of hostility toward foreigners, where democratic inclusion of outgroups is perceived as a threat to one's emerging ingroup identity. Moreover, the psychological distancing from mainstream society can increase openness to political violence, especially when such acts are framed as expressions of group-based resistance (Abrams & Hogg, 2006). Lastly, insights from social dominance theory (Sidanius & Pratto, 1999) suggest that norm-deviating individuals may be more inclined toward hierarchy-enhancing ideologies when they perceive existing norms as favoring egalitarianism or minority protection. In this framework, norm deviation can reflect a broader disposition to challenge status-quo values in favor of dominance-based ideologies. As such, norm deviation is expected to be positively associated with hostility toward foreigners, conspiracy mentality, and particularly with political affinity for violence to restore perceived group superiority or autonomy.

1.2.2. Intolerance and Political Polarization

Intolerance, defined as the unwillingness to permit disapproved beliefs or practices, is a central factor in political polarization. Unlike tolerance, which incorporates respect, freedom, and coexistence even amidst disagreement, intolerance rejects these principles, undermining the normative frameworks essential for social cohesion (Forst, 2013; Verkuyten et al., 2020, 2023). It manifests as exclusionary attitudes or behaviors that directly challenge pluralistic values and inclusivity. Adolescence, a critical stage for identity formation and social integration, is particularly susceptible to intolerance due to heightened sensitivity to peer influence and group dynamics.

Building again on social identity theory (Tajfel & Turner, 1979), intolerance strengthens ingroup cohesion by devaluing outgroups, especially when societal diversity is perceived as a threat to one's cultural or ideological identity (Verkuyten & Yogeeswaran, 2020). Adolescents with high intolerance may view inclusive democratic norms as illegitimately privileging "the other." This perceived asymmetry can foster rejection of democratic values and human rights discourses, especially when these are seen as institutional protections for outgroups. Reactance theory (Brehm, 1966) further explains how intolerance fuels resistance to perceived liberal or pluralistic societal expectations. When pluralistic norms (e.g., tolerance, inclusion) are interpreted as externally imposed constraints on group identity or cultural homogeneity, intolerant adolescents may react defensively—rejecting human rights frameworks as alien or overly permissive and developing conspiratorial ideation as a psychological response to perceived control (Sittenthaler et al., 2015). The perception that elites or institutions favor foreign or deviant groups can reinforce conspiracy thinking and ultimately justify political violence to restore perceived balance or justice. Within social dominance theory (Sidanius & Pratto, 1999), intolerance reflects a preference for group-based hierarchies and the legitimation of outgroup subordination. Adolescents high in intolerance may not only perceive other groups as threatening but also support ideologies that rationalize dominance and exclusion (Kteily et al., 2016). These attitudes are particularly predictive of hostility toward foreigners, conspiracy beliefs that protect dominant-group interests, and justification of violence against outgroups or institutions seen as enabling them.

1.2.3. Norm Deviation, Intolerance, and Political Polarization Across Gender, Age, and Migration Background

The relationships between norm deviation, intolerance, and the five indicators of political polarization are likely to vary across different demographic groups, including gender, age, and migration background. These factors influence the ways in which adolescents process social information, form identities, and engage with political ideologies, thereby shaping the nature and strength of these relationships.

Gender plays a significant role in how norm deviation and intolerance relate to political polarization. Research suggests that males are more likely to engage in oppositional or aggressive (political) behaviors, potentially due to socialization processes that emphasize competition, dominance, and outgroup hostility. For males, norm deviation may be more strongly linked to support for political violence and conspiracy thinking, while females may be more likely to experience norm deviation as a form of social exclusion, which may manifest as a rejection of democratic principles and human rights rather than aggression (Björkquist, 2018; Kuhn, 2010). Gender differences in political socialization may also influence the types of ingroup/outgroup dynamics that are reinforced through intolerance, with males more prone to adopting hierarchical, exclusionary ideologies (Sidanius et al., 1994).

Age is also a key variable that may moderate the impact of norm deviation and intolerance on political polarization. Adolescence is a period of heightened identity exploration and social influence (Eckstein et al., 2012; Quintelier, 2015), which makes younger individuals particularly susceptible to the effects of norm deviation. For younger adolescents, norm deviation may be more strongly associated with a rejection of democratic values and the embrace of extremist ideologies, as they may be more inclined to adopt rebellious or oppositional stances against perceived societal constraints. Older adolescents, on the other hand, may exhibit more refined ideological positions, with norm deviation potentially linking more to conspiracy theories and hostility toward foreigners, reflecting an increased awareness of societal power structures and their perceived exclusion from these structures (Wegemer, 2021).

Migration background is another crucial factor that may alter the relationship between norm deviation, intolerance, and political polarization. Adolescents with a migration background often navigate multiple cultural identities and may experience different forms of social marginalization and exclusion compared to their native peers (Sam & Berry, 2016). Studies show that migrant adolescents, especially those with a recent migration background, are more likely to experience alienation and reject mainstream democratic values, as these may be perceived as foreign to them or as instruments of social control used by the dominant group (Verkuyten, 2004). In contrast, native adolescents may be more prone to intolerance toward outgroups, as they are more likely to perceive threats to the status quo and to embrace exclusionary ideologies in response to perceived cultural dilution (Riek et al., 2006). For these adolescents, intolerance may more directly fuel hostility toward foreigners and the justification of political violence, whereas for migrant adolescents, norm deviation may be more strongly linked to the rejection of human rights and leaning toward conspiracy theories.

1.3. The Current Research

Norm deviation and intolerance are key psychological variables shaping political attitudes during adolescence—a phase of identity formation, autonomy seeking, and reevaluation of societal norms (Eckstein et al., 2012; Quintelier, 2015). Drawing on theories such as (a) reactance, (b) social identity, and (c) social dominance, this study examines how these tendencies relate to five indicators of political polarization: rejection of democracy, rejection of human rights, conspiracy mentality, hostility toward foreigners, and political affinity for violence.

Norm-deviating adolescents, as those opposing societal rules and norms, may reject democratic values and human rights as external constraints, triggering reactance and oppositional attitudes. They may also be more prone to conspiracy thinking, hostility toward foreigners, and justification of violence, especially when political discourses are seen as threatening or illegitimate. Intolerance, which is defined as rejection of differing beliefs and identities, can similarly foster political polarization. Intolerant adolescents may view democracy and human rights as overly permissive or protective of groups they perceive as threatening. This disposition is also linked to conspiracy beliefs, outgroup hostility, and legitimation of political violence in defense of perceived ingroup interests.

Despite clear theoretical assumptions, empirical studies on how norm deviation and intolerance contribute to adolescent political polarization remain scarce, even though adolescence is a key period for political socialization (Sears & Brown, 2013).

Demographic and contextual factors such as gender, age, and migration background may moderate these effects. Migrant adolescents may experience exclusion and alienation, potentially fostering alternative radicalization paths (Verkuyten, 2004), while native adolescents may react with stronger hostility toward foreigners. Gender differences also matter: Boys may lean toward aggression and political violence, while girls may express political disillusionment more passively (Björkquist, 2018; Kuhn, 2010). In the present study, we therefore examine the following hypotheses:

1. Adolescents who exhibit greater norm deviation and intolerance will demonstrate stronger tendencies to reject democracy and human rights, embrace conspiracy theories, express hostility toward foreigners, and show a higher affinity for political violence.
2. There will be significant differences in the relationships between norm deviation, intolerance, and the five indicators of political polarization based on gender, age, and migration background.

2. Method

2.1. Participants

The data for this study were derived from the Communities That Care (CTC) project in Lower Saxony (Germany), which is designed to assess risk and protective factors influencing youth behavior to inform evidence-based local prevention strategies (Groeger-Roth et al., 2018). Originally developed in the United States (Fagan et al., 2019; Hawkins & Catalano, 1992), the CTC framework has been adapted for use in Germany, aiming to empower communities to promote youth well-being and safety. The 2021/2022 survey (Soellner et al., 2022) included students from various school types across grades 6 to 11, as well as those attending vocational schools. Data collection employed a stratified random sampling approach and was conducted between November 29, 2021, and January 28, 2022.

In total, 316 schools, representing 373 classes and a total of 8,945 students, were contacted for participation in the CTC Youth Survey 2021/2022. Ultimately, 132 school classes from Lower Saxony took part, yielding a recruitment rate of 21.8% at the student level ($N = 1,948$). For this study, the sample was restricted to students in grade 8 and above, including equivalent grades in vocational school. This decision was grounded in developmental considerations (Eckstein & Noack, 2018; Sears & Brown, 2013), as it was assumed that the cognitive and ideological maturity required for valid engagement with political and ideological attitudes is more likely to be present in older adolescents. The resulting subsample consisted of $n = 1,211$, and all analyses were conducted using this subsample.

The mean age of participants was $M = 16.6$ years ($SD = 1.5$ years), with ages ranging from 15 to 22 years. Although our study focuses on adolescents, a small proportion of participants ($n = 102$; 8%) were 18 years or older. We retained these individuals in the analyses based on both theoretical and empirical considerations. Political socialization and the development of ideological attitudes do not end abruptly at age 18 but rather evolve gradually into early adulthood (Arnett, 2000). Including participants at the cusp of adulthood allows for a more comprehensive assessment of political attitudes across different youth age groups and ensures consistency with prior research that conceptualizes late adolescence and emerging adulthood as overlapping phases in the formation of political identity (Sears & Levy, 2003; Wray-Lake et al., 2020). Gender distribution was as follows: 45% identified as male, 53% as female, and 2% as diverse. Most

students were enrolled in grammar schools (gymnasium), vocational schools, or secondary schools. Approximately 30% of the respondents reported having a migration background, defined as either the adolescents themselves or both parents having been born outside of Germany. Notably, over 90% of these students were born in Germany. Detailed sample characteristics are presented in Table 1.

Table 1. Demographic characteristics of the sample.

| Variable | Categories | Proportion of students | |
		n	%
Gender	Male	545	45
	Female	642	53
	Diverse	24	2
Age*	15	296	24
	16	263	22
	17	345	28
	18	205	17
	19	64	5
	≥ 20	38	3
Type of school	Lower secondary school	15	1
	Intermediate secondary school	86	7
	German grammar school (gymnasium)	478	39
	Comprehensive school	330	28
	Vocational school	302	25
Migration background	No	848	70
	Yes	363	30

Notes: n = 1,211; * time of survey: November 2021 to January 2022.

2.2. Materials

The students received the questionnaire online and completed it during class whenever feasible under the conditions of the Covid-19 pandemic. In cases where this was not possible, students were allowed to complete the survey from home. The questionnaire comprised the full CTC Youth Survey 2021–2022, which covers a broad range of psychosocial and behavioral constructs. For the purposes of the present study, we selected and analyzed specific scales from this survey that captured the independent variables of norm deviation and intolerance, as well as indicators for polarized political attitudes and behavioral intentions. Each construct was measured using multiple items and subsequently aggregated into mean scales. The following paragraphs detail the individual measurement instruments (all items are listed in Supplementary File 2).

Norm deviation was assessed using a scale developed by Soellner et al. (2018), frequently employed in CTC research projects. The scale included three statements assessing individuals' propensity to deviate from general social norms and values. For example, students were asked to rate the acceptability of the statement "I do not abide by rules I do not like" on a four-point Likert scale (1 = *No* to 4 = *Yes*). The items were combined into a scale with an internal consistency of $\alpha = .66$ in the current sample. Higher scores reflected greater acceptance of norm-deviating behaviors among adolescents.

Intolerance was measured using items adapted from Beelmann and Karing (2015). Intolerance was conceptualized as the counterpart to tolerance, based on prior findings that negative evaluations tend to exhibit higher reliability than positive ones due to the reduced influence of social desirability effects (Beelmann et al., 2010). Intolerance was defined as an attitude pattern characterized by a minimal acceptance of deviations from social normative behavioral standards. Four items were rated on a five-point Likert scale (1 = *strongly disagree* to 5 = *strongly agree*). For example, participants evaluated the statement "If someone behaves differently than most people, I do not think it's good." The aggregated scale had an internal consistency of $\alpha = .73$, with higher values indicating greater intolerance.

Rejection of democracy was assessed through three items measuring fundamental opposition to democracy as a form of government and its principles. Participants rated statements such as "Democracy is the best way to govern a country" on a five-point Likert scale (1 = *strongly disagree* to 5 = *strongly agree*). The scale yielded an internal consistency of $\alpha = .51$, with higher values indicating stronger rejection of democracy.

Rejection of human rights was measured using two items generated and piloted by the research team. These items assessed a negative stance toward universal human rights. Participants rated the statements "All people are equal before the law" and "All people have freedom of religion" on a five-point Likert scale (1 = *strongly disagree* to 5 = *strongly agree*). The scale yielded an internal consistency of $\alpha = .54$, with higher scores indicating stronger rejection of human rights.

Conspiracy mentality was evaluated using two items from the General Conspiracist Beliefs Scale (Brotherton et al., 2013). Participants rated the statements "The government is hiding the truth from the public" and "The spread of certain diseases and viruses is in the interest of certain groups and organizations" on a five-point Likert scale (1 = *does not apply at all* to 5 = *applies completely*). These items were combined into a single scale with an internal consistency of $\alpha = .74$. Higher values indicating a stronger inclination toward conspiratorial thinking.

Hostility toward foreigners was assessed using three items derived from established surveys such as the Thuringia Monitor (e.g., Reiser et al., 2021) and the Authoritarianism Studies (e.g., Decker & Brähler, 2002). Participants rated statements such as "Foreigners are taking away our jobs" on a five-point Likert scale (1 = *does not apply at all* to 5 = *applies completely*). The aggregated scale demonstrated an internal consistency of $\alpha = .81$, with higher values reflecting greater hostility toward foreigners.

Affinity for political violence was measured as a behavioral indicator of polarized political intentions. This variable was operationalized using the violence-endorsing subscale of the Activism and Radicalism Intention Scale (Moskalenko & McCauley, 2009). Two items, one of which was "I would support an organization that fights for my group's political and legal rights even if the organization sometimes resorts to violence," were rated on a five-point Likert scale (1 = *does not apply at all* to 5 = *applies completely*). The resulting scale had an internal consistency of $\alpha = .84$, with higher scores indicating greater endorsement of violent actions to achieve political objectives.

3. Results

We used structural equation modeling (SEM) in JASP (JASP Team, 2024), employing R, to test our theoretical model with norm deviation and intolerance as latent predictors and rejection of democracy, rejection of human rights, conspiracy mentality, hostility toward foreigners, and affinity for political violence as latent outcomes. The overall model fit was evaluated using the χ^2/df ratio, CFI (comparative fit index), and RMSEA (root mean square error of approximation).

Descriptive statistics and correlations between the latent variables are presented in Table 2. Norm deviation was positively correlated with intolerance, rejection of democracy, rejection of human rights, conspiracy mentality, hostility toward foreigners, and affinity for political violence. Similarly, intolerance showed significant positive intercorrelations with all five indicators of polarized political attitudes and behavioral intentions.

Table 2. Correlations between the study variables and descriptive statistics.

Variable	M (SD)	2.	3.	4.	5.	6.	7.
1. Norm deviation	2.20 (0.62)	.25**	.08*	.09**	.17**	.13**	.10**
2. Intolerance	1.92 (0.73)	—	.14**	.22**	.10**	.27**	.13**
3. Rejection of democracy	2.34 (0.70)		—	.27**	.19**	.23**	.19**
4. Rejection of human rights	1.31 (0.65)			—	.11**	.37**	.11**
5. Conspiracy mentality	2.39 (1.09)				—	.29**	.22**
6. Hostility toward foreigners	1.67 (0.84)					—	.19**
7. Affinity for political violence	1.97 (1.12)						—

Notes: Norm deviation was measured on a four-point scale and the remaining variables on a five-point rating scale; * $p < .05$, ** $p < .01$.

To test for hypothesis 1, we first looked at the fit indices of the empirical model. The model fit was good, with fit values below or at recommended thresholds (χ^2(389.47)/df (131) = 2.97, CFI = .99, RMSEA = .03). The overall model is depicted in Figure 1. Norm deviation and intolerance were significantly interrelated (β = .18, $p < .01$). Norm deviation showed significant positive associations with rejection of democracy (β = .07, $p < .05$), conspiracy mentality (β = .21, $p < .01$), and affinity for political violence (β = .11, $p < .01$). However, no significant relationships were found between norm deviation and rejection of human rights (β = .02, $p < .61$) or hostility toward foreigners (β = .01, $p < .71$). Intolerance had significant positive associations with rejection of democracy (β = .14, $p < .01$), rejection of human rights (β = .30, $p < .01$), conspiracy mentality (β = .12, $p < .01$), hostility toward foreigners (β = .38, $p < .01$), and affinity for political violence (β = .08, $p < .01$). The model explained R^2 = .04 of the variances in rejection of democracy, R^2 = .09 in rejection of human rights, R^2 = .06 in conspiracy mentality, R^2 = .15 in hostility toward foreigners, and R^2 = .04 in affinity for political violence.
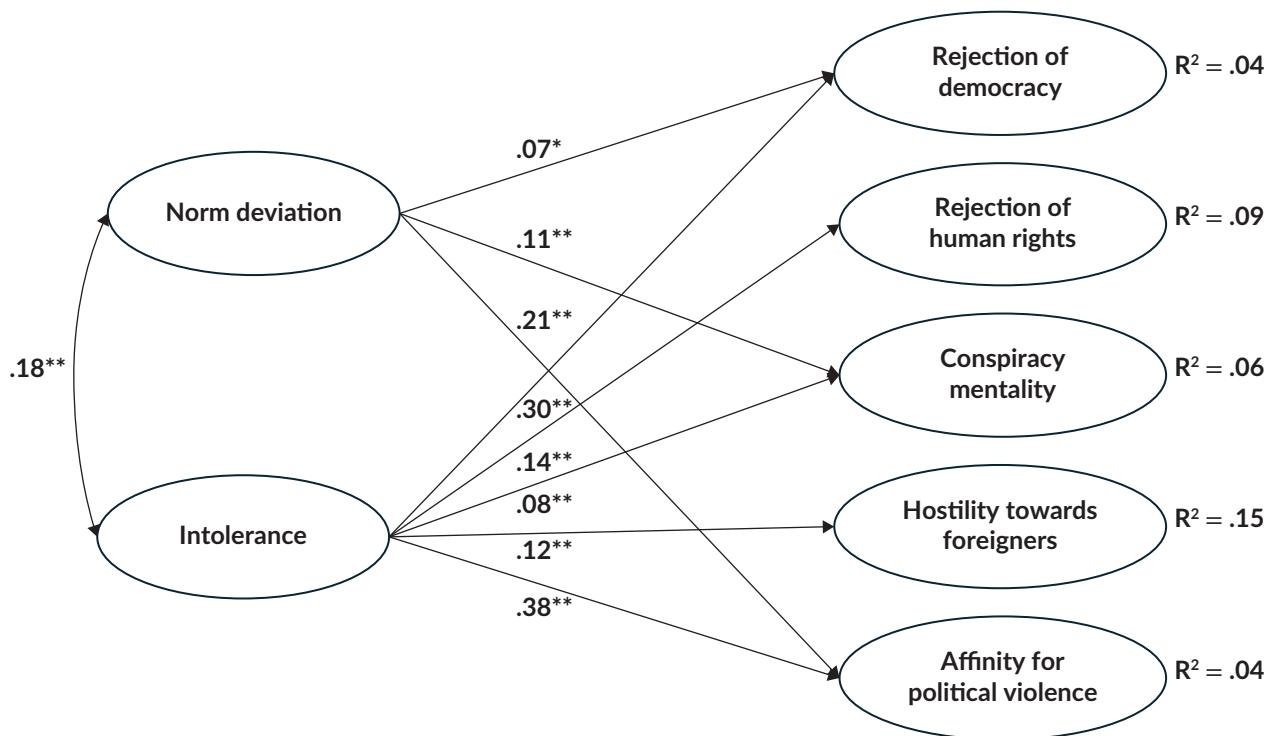
Figure 1. Specified SEM model with latent variables ($N = 1{,}211$, $\chi^2(389.47)/df\ (131) = 2.97$, CFI $= .99$, RMSEA $= .03$). Notes: Path coefficients represent standardized estimates; associations between norm deviation and indicators of political polarization are above the arrows, while associations between intolerance and indicators of political polarization are shown below the arrows; only significant paths are included; ** $p < .01$, * $p < .05$.

In line with hypothesis 2, we also examined the relationships between predictors and outcomes across different demographic groups, considering, gender, age, and migration background. The means of norm deviation ($M_{male} = 2.27$ vs. $M_{female} = 2.13$, $F\ [df = 1, 1210] = 14.88$, $p < .001$) and intolerance ($M_{male} = 2.05$ vs. $M_{female} = 1.81$, $F\ [df = 1, 1210] = 32.69$, $p < .001$) differed significantly by gender but not by age or migration background. Table 3 summarizes the main results of the applied multi-group SEM. Detailed information on the multi-group SEM models for gender (Table S1), age (Table S2), and migration background (Table S3) is provided in Supplementary File 1.

First, the multi-group SEM revealed distinct patterns in path coefficients based on the reported gender of participants (students identifying as "diverse" were excluded, $n = 24$). Two main patterns emerged for norm deviation:

> Asymmetric relationships: Norm deviation was more strongly and positively associated with the rejection of democracy and human rights among male participants, but negatively among female participants.

> Symmetric relationships: Positive associations with conspiracy mentality and affinity for political violence were observed across genders but were slightly more pronounced among males in case of political violence.

Table 3. Path differences in multi-group SEM models between norm deviation and intolerance and indicators of polarized political attitudes and behavioral intentions by gender, age, and migration background.

Polarization indicators	Norm deviation		Intolerance	
	Groups	z	Groups	z
Rejection of democracy	f < m	3.42**	f > m	2.17*
	y16 > o16	0.29	y16 = o16	0.18
	no MB < MB	0.77	no MB = MB	0.50
Rejection of human rights	f < m	3.38**	f > m	4.98**
	y16 < o16	3.50**	y16 > o16	3.79**
	no MB < MB	2.68*	no MB > MB	6.41**
Conspiracy mentality	f = m	1.05	f = m	0.31
	y16 > o16	1.82†	y16 > o16	1.05
	no MB > MB	2.16*	no MB > MB	0.82
Hostility toward foreigners	f = m	0.43	f > m	1.47†
	y16 = o16	1.26	y16 = o16	0.20
	no MB < MB	2.04*	no MB > MB	3.47**
Affinity for political violence	f < m	1.47†	f = m	0.03
	y16 = o16	0.20	y16 > o16	2.48*
	no MB < MB	0.37	no MB > MB	3.50**

Notes: f = female, m = male, y16 = younger than 16 years, o16 = older than 16 years, no MB = no migration background, MB = migration background; z-values indicate whether the strength of paths differs significantly between groups; † $p < .10$, * $p < .05$, ** $p < .01$.

For intolerance, positive associations with rejection of democracy, rejection of human rights, and hostility toward foreigners were significant for both genders but more pronounced among females. The relationships between intolerance and conspiracy mentality, as well as affinity for political violence, were positive and comparable across genders.

Second, we analyzed the relationships between norm deviation and intolerance as a function of age. Participants were divided into two groups: younger than 16 years and 16 years or older. For norm deviation, positive associations with rejection of democracy, conspiracy mentality, and affinity for political violence were significant among younger participants. Conversely, the relationship between norm deviation and rejection of human rights was more pronounced among older participants. Hostility toward foreigners showed no significant association in either age group. For intolerance, relationships with rejection of democracy and hostility toward foreigners were similarly strong across both age groups. However, stronger associations emerged between intolerance and rejection of human rights and affinity for political violence among younger participants. Furthermore, the relationship between intolerance and conspiracy mentality was positive and significant only in case of younger participants.

Third, we analyzed the relationships between norm deviation and intolerance and the five indicators of polarized political attitudes and behavioral intentions, differentiating the sample based on participants' migration background. For norm deviation, positive associations with rejection of human rights and hostility toward foreigners were stronger among participants with a migration background. In addition, the relationship between norm deviation and rejection of democracy and affinity for political violence was

only significant for students reporting a migration background. Conversely, participants without a migration background showed a stronger association between norm deviation and conspiracy mentality. For intolerance, the relationships with rejection of human rights, hostility toward foreigners, and affinity for political violence were stronger among participants without a migration background. In addition, the association between intolerance and conspiracy mentality was significant and positive only for students without a migration background. The relationship between intolerance and rejection of democracy was similarly strong across groups.

4. Discussion

This study aimed to examine how norm deviation and intolerance relate to key indicators of political polarization among adolescents. Additionally, it explored demographic differences in the relationships of norm deviation and intolerance with rejection of democracy and human rights, conspiracy mentality, hostility toward foreigners, and affinity for political violence by gender, age, and migration background. We found that norm deviation and intolerance are significant predictors of polarized political attitudes and behavioral intentions, albeit with varying degrees of influence across different outcomes and demographic groups.

The findings largely support the assumptions of hypothesis 1 and offer deeper insight into the differential predictive value of norm deviation and intolerance in explaining indicators of political polarization among adolescents. As anticipated, intolerance proved to be associated with all five indicators. This is in line with the conceptualization of intolerance as a broad social-psychological factor that reflects a negative orientation toward diversity and perceived norm deviations of others (Verkuyten et al., 2020, 2023; Whitley & Webster, 2018). In contrast, norm deviation—understood here as adolescents' tendency to distance themselves from established societal norms and rule-breaking behavior—exhibited a more selective pattern of associations. It was related to the rejection of democracy, conspiracy mentality, and political affinity for violence, but showed no significant associations with the rejection of human rights or hostility toward foreigners. This pattern highlights the conceptual distinction between norm deviation as a behavioral style of rule-breaking and autonomy assertion and intolerance as a value-laden, ideologically charged orientation toward others. The partial alignment of norm deviation with elements of reactance theory (Miron & Brehm, 2006) and social identity theory (Tajfel & Turner, 1979) suggests that adolescent rule-breaking may foster political alienation and mistrust, particularly toward institutional authority. However, the absence of associations with human rights rejection or outgroup derogation suggests that such deviance does not necessarily translate into ideologically structured attitudes. Furthermore, the lack of a robust link to hostility toward foreigners challenges assumptions drawn from social dominance theory (Sidanius & Pratto, 1999), indicating that norm-deviating adolescents do not necessarily endorse hierarchy-enhancing ideologies. Instead, norm deviation may reflect diffuse opposition, psychological disengagement, or situational frustration rather than structured ideological positioning.

The demographic differences added depth to the findings and predominantly supported hypothesis 2. Gender-specific patterns showed that norm deviation was more strongly linked to rejection of democracy and human rights among males, while intolerance had greater predictive power among females. These asymmetries likely reflected gendered socialization—males are often socialized toward autonomy and resistance (Hogg & Vaughan, 2020), whereas females tend to be more sensitive to relational dynamics (Eagly & Wood, 2012), which may increase susceptibility to exclusionary attitudes. Age also played a

significant role: Younger adolescents showed stronger links between norm deviation and democratic rejection, conspiracy mentality, and political violence, consistent with early adolescence as a period of identity exploration and emotional reactivity (Crone & Dahl, 2012). In contrast, older adolescents exhibited stronger associations between intolerance and exclusionary attitudes, likely reflecting more advanced ideological development and increasing sensitivity to socio-political cues (Abrams & Hogg, 2006). Migration background revealed further nuance. Among migrant adolescents, norm deviation was more strongly associated with rejecting human rights and endorsing violence—potentially due to experiences of exclusion and marginalization (Verkuyten, 2004). For non-migrant youth, stronger links between norm deviation and conspiracy mentality may have reflected perceived cultural threat and identity insecurity, aligning with integrated threat theory (Stephan & Stephan, 2000). These findings underscore how gender, age, and migration background shape distinct pathways to political polarization.

Overall, the findings highlight intolerance as a stable predictor of different indicators of political polarization, while norm deviation emerges as a more situational and demographically shaped pathway. This underscores the need for a developmental and context-sensitive lens on how adolescents relate to democracy and social cohesion.

4.1. Limitations

The present study, while offering valuable insights into the psychological roots of polarized political attitudes among adolescents, is not without limitations that warrant discussion.

First, the theoretical model was assessed using a limited number of items for each construct, a constraint imposed by the broad scope of the CTC study (Reder et al., 2024), which included numerous other constructs such as substance use, depressive symptomatology, and experiences during the Covid-19 pandemic. To mitigate potential validity concerns, the selected items were informed by prior research and optimized for conceptual coverage within these constraints. Future studies should adopt a more comprehensive approach to construct measurement, as this would enable a more nuanced understanding of the interplay between psychological factors and political polarization.

Second, the wording of some items may have introduced comprehension challenges, particularly for younger participants. While the items were designed to align with the reading level of adolescents, their complexity may have exceeded the cognitive capacities of younger school children. In subsequent research, simpler phrasing and additional pretesting with age-specific samples could enhance accessibility and validity.

Third, conducting the study during the Covid-19 pandemic presented unique limitations associated with online data collection. For participants under 18, parental consent was required, introducing a potential selection bias if parents restricted participation. Additionally, online studies carry inherent risks, such as uncertain respondent authenticity, which have been documented in prior research (e.g., Newman et al., 2020). To counter these issues and enhance the validity of findings, the study employed representative sampling for Lower Saxony and weighting of relevant demographic indicators (Soellner et al., 2022).

Participants completed the online survey either at home or in school, which could have been influenced by external factors like the presence of parents or teachers. Although the survey was anonymous and participants

were instructed to respond independently, the setting may have impacted their openness to sensitive topics such as norm deviation or attitudes toward democracy. Most data (99%) were collected in school settings, where teachers were asked not to intervene. Future research should better control for these contextual factors to enhance the ecological validity of adolescent political attitude studies.

The sample's age distribution presents another limitation, as it was heavily centered around 16 years, with insufficient representation of older adolescents (18–22 years) to establish a separate age group for nuanced analyses. Although our primary focus lay on adolescents, we included this small subgroup of participants aged 18 years and older in our analyses. This decision was theoretically grounded in research on emerging adulthood—which conceptualizes late adolescence and early adulthood as a continuous developmental stage in the formation of political beliefs (Arnett, 2000).

A conceptual limitation concerns our use of the term *political polarization*. While polarization typically refers to a growing divide between social groups, our study does not track such dynamics directly. Instead, we examine individual attitudes—such as rejection of democratic values or affinity for political violence—that are widely seen as early indicators or precursors of polarization. In this broader perspective, our use aligns with developmental and political psychology approaches that explore early risk factors for societal fragmentation during adolescence.

Our study provides valuable insights into norm deviation and political attitudes, but its findings may not be easily generalized to different cultural or political contexts. Norm deviation, particularly in relation to democracy and human rights, may have different meanings in non-WEIRD (Western, Educated, Industrialized, Rich, Democratic) societies. While the behaviors linked to norm deviation may occur in other democratic contexts, their expression could vary based on political and cultural factors. Future research should explore how norm deviation manifests in different political systems to better understand its relationship with polarized political attitudes across diverse cultures.

Lastly, the measure of hostility toward foreigners did not include parallel items assessing prejudice or hostility toward Germans from the perspective of participants with migration backgrounds. This asymmetry may have limited the scope of intergroup analyses, as the dynamics of prejudice and hostility often manifest reciprocally and are shaped by both majority and minority group perspectives (Lutterbach & Beelmann, 2023; Raabe & Beelmann, 2011).

4.2. Implications for Future Research and Prevention

Future research should investigate how identity development, socialization, and peer dynamics explain demographic differences in polarization. Stronger associations among younger and migrant adolescents suggest that developmental and cultural contexts interact in shaping political attitudes (Crocetti et al., 2023; Sam & Berry, 2016). Including variables such as identity conflict or bicultural stress could clarify specific vulnerabilities. Examining older adolescents (18–22) is also crucial, as this phase marks a key period for political and identity consolidation (Sears & Brown, 2013; Zaff et al., 2011).

The gender differences identified—especially regarding the link between norm deviation and rejection of democratic values—highlight the need to explore how gendered roles and expectations shape pathways to

political polarization. Examining gender-specific socialization and identity formation could inform more targeted interventions (Kish Bar-On et al., 2024; Paluck & Green, 2009). The absence of strong associations between norm deviation and hostility toward foreigners across age and gender groups points to the need for more nuanced measures of intergroup hostility. These should reflect a wider range of target groups and perspectives, particularly those of individuals with migration backgrounds, to better capture the complexity of intergroup dynamics in diverse societies. Including concepts like bicultural or hybrid identities (Schwartz et al., 2017) may further clarify how identity complexity influences the relationship between norm deviation, intolerance, and polarized political attitudes.

The findings of this study have important implications for prevention programs. For younger adolescents, civic education should emphasize democratic values, critical thinking, and media literacy to reduce susceptibility to conspiracy beliefs and political violence. Among adolescents with migration backgrounds, interventions should address acculturation stressors and promote inclusive identity formation to mitigate the effects of norm deviation on polarized attitudes. Gender-sensitive approaches are also essential. Programs for male adolescents should encourage positive peer influences and provide prosocial role models to counter democratic rejection and political violence. For female adolescents, interventions should focus on reducing intolerance-related risks—such as rejection of human rights or intergroup hostility—by fostering empathy, dialogue, and social cohesion.

Strong links between intolerance and indicators like conspiracy mentality and political violence underscore the urgency of comprehensive civic and media literacy programs (Goren & Yemini, 2017; Schmitt et al., 2018). These should empower adolescents to critically engage with polarized narratives, especially online. Community-based initiatives promoting shared projects and shared realities (Lutterbach & Beelmann, 2020) can strengthen mutual understanding. Additionally, fostering inclusive norms—centered on shared values, diversity, and equal participation—can help prevent polarization and support social cohesion (Shani et al., 2023). Addressing these challenges can ultimately contribute to a more inclusive and democratic society.

5. Conclusion

This study contributes to the discourse on political polarization by showing how norm deviation and intolerance relate to polarized attitudes among adolescents, with clear differences across gender, age, and migration background. These patterns highlight the need to understand the developmental and social-psychological roots of polarization (Beelmann & Lutterbach, 2022). Effective prevention must reflect these demographic nuances: Early adolescence is a key window for fostering democratic values and critical thinking, gender-sensitive strategies should address distinct social experiences, and tailored acculturation efforts can support integration and reduce polarization among youth with and without migration backgrounds. Targeting these factors may help prevent intolerance and societal division during times of political and cultural transformation.

References

Abrams, D., & Hogg, M. A. (2006). *Social identifications: A social psychology of intergroup relations and group processes*. Routledge.

Arnett, J. J. (2000). Emerging adulthood: A theory of development from late teens through the twenties. *American Psychologist*, *55*(5), 469–480. https://doi.org/10.1037/0003-066X.55.5.469

Bakker, B. N., & Lelkes, Y. (2024). Putting the affect into affective polarisation. *Cognition and Emotion*, *38*(4), 418–436. https://doi.org/10.1080/02699931.2024.2362366

Beelmann, A. (2020). A social-developmental model of radicalization: A systematic integration of existing theories and empirical research. *International Journal of Conflict and Violence, 14*(1). https://doi.org/10.4119/ijcv-3778

Beelmann, A. (2021). Concept of and approaches toward a developmental prevention of radicalization: Promising strategies to keep young people away from political, religious and other forms of extremism. *Monatsschrift für Kriminologie und Strafrechtsreform/Journal of Criminology and Penalty Reform*, *104*(3), 298–309. https://doi.org/10.1515/mks-2021-0130

Beelmann, A., & Karing, C. (2015). Förderung toleranter Einstellungen und die Prävention von Vorurteilen. Langzeitwirkungen des Programms zur Förderung von Akzeptanz, Respekt, Toleranz und sozialer Kompetenz (PARTS). *Forum Kriminalprävention*, *1*, 51–58.

Beelmann, A., & Lutterbach, S. (2022). Developmental prevention of prejudice: Conceptual issues, evidence-based designing, and outcome results. *Review of General Psychology*, *26*(3), 298–316. https://doi.org/10.1177/10892680211056314

Beelmann, A., & Lutterbach, S. (2023). Special issue editorial: Political socialization as desiderata of developmental psychology. *International Journal of Developmental Science*, *17*(1/2/3), 3–5. https://doi.org/10.3233/DEV-230351

Beelmann, A., Saur, M., & Ziegler, P. (2010). *Thüringer Studie zur Vorurteilsprävention und Toleranzentwicklung. Projektbericht*. Friedrich Schiller University Jena, Institute of Psychology.

Berg-Schlosser, D., Badie, B., & Morlino, L. (Eds.). (2020). *The Sage handbook of political science*. Sage.

Björkquist, K. (2018). Gender differences in aggression. *Current Opinion in Psychology*, *19*, 39–42. https://doi.org/10.1016/j.copsyc.2017.03.030

Bliuc, A.-M., Betts, J. M., Vergani, M., Bouguettaya, A., & Cristea, M. (2024). A theoretical framework for polarization as the gradual fragmentation of a divided society. *Communications Psychology*, *2*, Article 75. https://doi.org/10.1038/s44271-024-00125-1

Brehm, J. W. (1966). *A theory of psychological reactance*. Academic Press.

Brotherton, R., French, C. C., & Pickering, A. D. (2013). Measuring belief in conspiracy theories: The generic conspiracist belief scale. *Frontiers in Psychology*, *4*, Article 00279. https://doi.org/10.3389/fpsyg.2013.00279

Crocetti, E., Albarello, F., Meeus, W., & Rubini, M. (2023). Identities: A developmental social-psychological perspective. *European Review of Social Psychology*, *34*(1), 161–201. https://doi.org/10.1080/10463283.2022.2104987

Crone, E., & Dahl, R. (2012). Understanding adolescence as a period of social-affective engagement and goal flexibility. *Nature Reviews Neuroscience*, *13*, 636–650. https://doi.org/10.1038/nrn3313

Decker, O., & Brähler, E. (Eds.). (2002). *Autoritäre Dynamiken in unsicheren Zeiten. Leipziger Autoritarismus Studie 2022*. Psychosozial-Verlag.

Dekker, H., German, D. B., & De Landtsheer, C. (2020). Political socialization theory, research, and application. History and analysis of forty years of the Research Committee on Political Socialization and Education of the International Political Science Association: 1997–2019. *Politics, Culture and Socialization*, *8*, 34–80.

Eagly, A. H., & Wood, W. (2012). Social role theory. In P. A. M. Van Lange, A. W. Kruglanski, & E. T. Higgins (Eds.), *Handbook of theories of social psychology* (pp. 458–476). Sage.

Eckstein, K., & Noack, P. (2018). Political socialization. In B. Gniewosz & P. F. Titzmann (Eds.), *Handbook of youth. Psychological perspectives on development in* adolescence (pp. 485–495). Kohlhammer.

Eckstein, K., Noack, P., & Gniewosz, B. (2012). Attitudes towards political engagement and willingness to participate in politics: Trajectories throughout adolescence. *Journal of Adolescence*, *35*, 485–495. https://doi.org/10.1016/j.adolescence.2011.07.002

Fagan, A. A., Bumbarger, B. K., Barth, R. P., Bradshaw, C. P., Rhoades Cooper, B., Supplee, L. H., & Klein Walker, D. (2019). Scaling up evidence-based interventions in US public systems to prevent behavioral health problems: Challenges and opportunities. *Prevention Science*, *20*, 1147–1168. https://doi.org/10.1007/s11121-019-01048-8

Forst, R. (2013). *Toleration in conflict: Past and present*. Cambridge University Press.

Goren, H., & Yemini, M. (2017). Global citizenship education redefined—A systematic review of empirical studies on global citizenship education. *International Journal of Educational Research*, *82*, 170–183. https://doi.org/10.1016/j.ijer.2017.02.004

Groeger-Roth, F., Marks, E., & Meyer, A. (2018). Erfahrungen des Landespräventionsrates Niedersachsen (LPR) bei der Umsetzung einer qualitätsorientierten und evidenzbasierten Präventionspraxis. In M. Walsh, B. Pniewski, M. Kober, & A. Armborst (Eds.), *Evidenzorientierte Kriminalprävention in Deutschland* (pp. 145–165). Springer.

Hawkins, J. D., & Catalano, R. F., Jr. (1992). *Communities that care: Action for drug abuse prevention*. Jossey-Bass.

Hewstone, M., Stroebe, W., & Jonas, K. (2021). *Social psychology*. Springer.

Hogg, M. A., & Vaughan, G. M. (2020). *Social psychology* (9th ed.). Pearson.

Imhoff, R., & Bruder, M. (2014). Speaking (un-)truth to power: Conspiracy mentality as a generalised political attitude. *European Journal of Personality*, *28*(1), 25–43. https://doi.org/10.1002/per.1930

JASP Team. (2024). *JASP* (Version 0.10.0) [Computer software].

Kish Bar-On, K., Diamant, E., Lelkes, Y., & Rand, D. G. (2024). Unraveling polarization: Insights into individual and collective dimensions. *PNAS Nexus*, *3*(10), Article pgae426. https://doi.org/10.1093/pnasnexus/pgae426

Kteily, N., Hodson, G., & Bruneau, E. (2016). They see us as less than human: Metadehumanization predicts conflict via reciprocal dehumanization. *Journal of Personality and Social Psychology*, *110*(3), 343–370. https://doi.org/10.1037/pspa0000044

Kuhn, H. (2010). International perspectives on political socialization and gender: An introduction. In A. Ittel, H. Merkens, L. Stecher, & J. Zinnecker (Eds.), *Yearbook of youth research* (pp. 11–24). VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-531-92320-8_1

Lelkes, Y. (2016). Mass polarization: Manifestations and measurements. *Public Opinion Quarterly*, *80*(S1), 392–410. https://doi.org/10.1093/poq/nfw005

Lutterbach, S., & Beelmann, A. (2020). Positive and negative intergroup contact and shared reality: Contact effects among host society and refugees. *European Journal of Social Psychology*, *50*(1), 143–159. https://doi.org/10.1002/ejsp.2601

Lutterbach, S., & Beelmann, A. (2023). Relations between positive and negative extended contact and prejudice in host society and refugees: Effects of positive and negative direct contact. *International Journal of Intercultural Relations*, *93*, Article 101760. https://doi.org/10.1016/j.ijintrel.2023.101760

Miron, A. M., & Brehm, J. W. (2006). Reactance theory–40 years later. *Zeitschrift für Sozialpsychologie*, *37*(1), 9–18. https://doi.org/10.1024/0044-3514.37.1.9

Moskalenko, S., & McCauley, C. (2009). Measuring political mobilization: The distinction between activism and radicalism. *Terrorism & Political Violence*, *21*(2), 239–260. https://doi.org/10.1080/09546550902765508

Neundorf, A., & Smets, K. (2017). Political socialization and the making of citizens. In *Oxford handbook topics in politics*. Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199935307.013.98

Newman, A., Bavik, Y. L., Mount, M., & Shao, B. (2020). Data collection via online platforms: Challenges and recommendations for future research. *Applied Psychology*, *70*(3), 1380–1402. https://doi.org/10.1111/apps.12302

Osborne, D., Costello, T. H., Duckitt, J., & Sibley, C. G. (2023). The psychological causes and societal consequences of authoritarianism. *Nature Reviews Psychology*, *2*, 220–232. https://doi.org/10.1038/s44159-023-00161-4

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. https://doi.org/10.1146/annurev.psych.60.110707.163607

Piazza, J. A. (2023). Political polarization and political violence. *Security Studies*, *32*(3), 476–504. https://doi.org/10.1080/09636412.2023.2225780

Quintelier, E. (2015). Engaging adolescents in politics: The longitudinal effect of political socialization agents. *Youth & Society*, *47*(1), 51–69. https://doi.org/10.1177/0044118X13507295

Raabe, T., & Beelmann, A. (2011). Development of ethnic, racial, and national prejudice in childhood and adolescence: A multinational meta-analysis of age differences. *Child Development*, *82*(6), 1715–1737. https://doi.org/10.1111/j.1467-8624.2011.01668.x

Reder, M., Runge, R. A., Schlüter, H., & Soellner, R. (2024). The German Communities That Care Youth Survey: Dimensionality and validity of risk factors. *Frontiers in Public Health*, *12*, Article 1472347. https://doi.org/10.3389/fpubh.2024.1472347

Reiser, M., Küppers, A., Hebenstreit, J., Salheiser, A., & Vogel, L. (2021). *Demokratie in der Corona-Pandemie: Ergebnisse des Thüringen Monitors 2021*. Friedrich Schiller University Jena, Center for Research on Right-Wing Extremism, Democracy Education, and Social Integration.

Rekker, R., Keijsers, L., Branje, S., & Meeus, W. (2015). Political attitudes in adolescence and emerging adulthood: Developmental changes in mean level, polarization, rank-order stability, and correlates. *Journal of Adolescence*, *41*(1), 136–147. https://doi.org/10.1016/j.adolescence.2015.03.011

Renström, E. A., Bäck, H., & Carroll, R. (2022). Protecting the ingroup? Authoritarianism, immigration attitudes, and affective polarization. *Frontiers in Political Science*, *4*, Article 919236. https://doi.org/10.3389/fpos.2022.919236

Riek, B. M., Mania, E. Q., & Gaertner, S. L. (2006). Intergroup threat and outgroup attitudes: A meta-analytic review. *Personality and Social Psychology Review*, *10*(4), 336–353. https://doi.org/10.1207/s15327957pspr1004_4

Sam, D. L., & Berry, J. W. (Eds.). (2016). *The Cambridge handbook of acculturation psychology* (2nd ed.). Cambridge University Press.

Schmitt, J. B., Rieger, D., Ernst, J., & Roth, H. (2018). Critical media literacy and Islamist online propaganda: The feasibility, applicability and impact of three learning arrangements. *International Journal of Conflict and Violence*, *12*, Article a642. https://doi.org/10.4119/ijcv-3104

Schwartz, S. J., Birman, D., Benet-Martínez, V., & Unger, J. B. (2017). Biculturalism: Negotiating multiple cultural streams. In S. J. Schwartz & J. B. Unger (Eds.), *The Oxford handbook of acculturation and health* (pp. 29–47). Oxford University Press.

Sears, D. O., & Brown, C. (2013). Childhood and adult political development. In L. Huddy, D. O. Sears, & J. S. Levy (Eds.), *The Oxford handbook of political psychology* (2nd ed., pp. 59–95). Oxford University Press.

Sears, D. O., & Levy, S. (2003). Childhood and adult political development. In D. O. Sears, L. Huddy, & R. Jervis (Eds.), *The Oxford handbook of political psychology* (1st ed., pp. 60–109). Oxford University Press.

Shani, M., de Lede, S., Richters, S., Kleuker, M., Middendorf, W., Liedtke, J., Witolla, S., & van Zalk, M. (2023). A social network intervention to improve adolescents' intergroup tolerance via norms of equality-based respect: The "Together for Tolerance" feasibility study. *International Journal of Developmental Science*, *17*(1/2/3), 93–110. https://doi.org/10.3233/DEV-230342

Sherif, M. (1966). *The psychology of social norms*. Harper Torchbooks.

Sidanius, J., & Pratto, F. (1999). *Social dominance: An intergroup theory of social hierarchy and oppression*. Cambridge University Press. https://doi.org/10.1017/CBO9781139175043

Sidanius, J., Pratto, F., & Bobo, L. (1994). Social dominance orientation and the political psychology of gender: A case of invariance? *Journal of Personality and Social Psychology*, *67*(6), 998–1011. https://doi.org/10.1037/0022-3514.67.6.998

Sittenthaler, S., Traut-Mattausch, E., & Jonas, E. (2015). Observing the restriction of another person: Vicarious reactance and the role of self-construal and culture. *Frontiers in Psychology*, *6*, Article 01052. https://doi.org/10.3389/fpsyg.2015.01052

Soellner, R., Reder, M., & Frisch, J. U. (2018). *Communities that care: Student survey in Lower Saxony 2017*. Hildesheim University Press.

Soellner, R., Schlüter, H., & Kollek, M. (2022). *Communities that care: Student survey in Lower Saxony 2021/2022*. Hildesheim University Press.

Stephan, W. G., & Stephan, C. W. (2000). An integrated threat theory of prejudice. In S. Oskamp (Ed.), *Reducing prejudice and discrimination* (pp. 23–45). Lawrence Erlbaum Associates.

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). Brooks/Cole.

Tyler, M., & Iyengar, S. (2023). Learning to dislike your opponents: Political socialization in the era of polarization. *American Political Science Review*, *117*(1), 347–354. https://doi.org/10.1017/S000305542200048X

Verkuyten, M. (2004). *The social psychology of ethnic identity*. Psychology Press.

Verkuyten, M., Adelman, L., & Yogeeswaran, K. (2020). The psychology of intolerance: Unpacking diverse understandings of intolerance. *Current Directions in Psychological Science*, *29*(5), 467–472. https://doi.org/10.1177/0963721420924763

Verkuyten, M., & Yogeeswaran, K. (2020). Cultural diversity and its implications for intergroup relations. *Current Opinion in Psychology*, *32*, 1–5. https://doi.org/10.1016/j.copsyc.2019.06.010

Verkuyten, M., Yogeeswaran, K., & Adelman, L. (2023). The social psychology of intergroup tolerance and intolerance. *European Review of Social Psychology*, *34*(1), 1–43. https://doi.org/10.1080/10463283.2022.2091326

Wegemer, C. M. (2021). Critical civic motivation of marginalized youth: An expectancy-value approach. *Youth & Society*, *55*(2), 213–237. https://doi.org/10.1177/0044118X211044522

Whitley, B. E., & Webster, G. D. (2018). The relationships of intergroup ideologies to ethnic prejudice: A meta-analysis. *Personality and Social Psychology Review*, *23*(3), 207–237. https://doi.org/10.1177/1088868318761423

Wray-Lake, L. (2019). How do young people become politically engaged? *Child Development Perspectives*, *13*(2), 127–132. https://doi.org/10.1111/cdep.12324

Wray-Lake, L., Arruda, E. H., & Schulenberg, J. E. (2020). Civic development across the transition to adulthood in a national U.S. sample: Variations by race/ethnicity, parent education, and gender. *Developmental Psychology*, *56*(10), 1948–1967. https://doi.org/10.1037/dev0001101

Zaff, J. F., Kawashima-Ginsberg, K., Lin, E. S., Lamb, M., Balsano, A., & Lerner, R. M. (2011). Developmental trajectories of civic engagement across adolescence: Disaggregation of an integrated construct. *Journal of Adolescence*, *34*(6), 1207–1220. https://doi.org/10.1016/j.adolescence.2011.07.005

About the Authors

Sebastian Lutterbach (PhD) is a research associate in the Department of Research Synthesis, Intervention and Evaluation at the Institute of Psychology, Friedrich Schiller University Jena, Germany. His research focuses on acculturation processes between majority and minority groups, intergroup contact, and the development and prevention of extremism, polarization, and prejudice.

Andreas Beelmann (PhD) is a professor and the head of the Department of Research Synthesis, Intervention and Evaluation, and a member of the Research Center on Right-Wing Extremism and Civic Education at Friedrich Schiller University Jena, Germany. His research focuses on the development and prevention of behavioral problems in childhood and adolescence.

cogitatio

ARTICLE

Associations Between Perceived Societal Polarisation and (Extreme) Non-Normative Attitudes and Behaviour

Rebekka Kesberg [1] , **Allard R. Feddes** [2] , **Eva Vogel** [2] , and **Bastiaan T. Rutjens** [2]

[1] School of Psychology, University of Sussex, UK
[2] Department of Social Psychology, University of Amsterdam, The Netherlands

Correspondence: Rebekka Kesberg (r.kesberg@sussex.ac.uk)

Abstract

Perceptions of societal polarisation in a country may influence individuals' willingness to engage in non(normative) collective action. In the present research, we test the hypothesis that perceived societal polarisation reduces trust in the government, particularly when the government is perceived as posing a threat to the own social group. In turn, we expect increased willingness to engage in collective action on behalf of the ingroup. To test our predictions, we used a 2 (no threat versus threat) × 2 (no polarisation versus polarisation) experimental design. In the paradigm, participants are citizens in a fictitious country called "Bovenland." Participants read three newspaper articles about political issues targeted at their ingroup "the Southerners." After each newspaper article, participants indicated their intention to engage in collective action as well as their trust in the Bovenland government. A pilot study ($N = 42$) suggested that our experimental manipulation of perceived threat and perception of polarisation was effective. Our (preregistered) main study ($N = 982$) gathered through Prolific, found that perceived threat, but not polarisation, resulted in significantly more collective action intentions (normative, non-normative, and extreme non-normative). Under threat (but not under polarisation) significantly more (non)normative action intentions emerged. Bootstrapping analyses supported our finding that there was no direct effect of polarisation on collective action intentions. However, in the no-threat condition, polarisation increased trust in the Bovenland government, which predicted less collective action intentions (normative, non-normative, and extreme non-normative). The implications of these findings will be discussed.

Keywords

collective action intentions; normative behaviour; perceived threat; polarisation; political trust

1. Introduction

In recent years, increased protests and other forms of collective action directed against governments and other authorities emerged in Western democracies. For example, the "yellow vests" protests emerged after proposed increases to gas prices and developed into a wider protest against the government (Breeden & Specia, 2019; Jetten et al., 2020). In other countries, both climate protesters (Schuetze, 2023) and farmers ("Boeren dumpen puin," 2024a) blocked highways as a means of protest against government actions. Other protests included violent elements, such as pro-Palestine protesters on campuses in the US and Europe ("Politie: niet-vreedzame," 2024b; "Scenes from the student protests," 2024) or far-right rioters in the UK attacking the police ("110 police officers injured," 2024). While these protests differ in their underlying motivation, they are a form of collective action in response to perceived threats. Specifically, a reaction to the feeling that their group is being disadvantaged by the government or that the government is unable to deal with a threat. This implies how people perceive the society and world they live in shapes their response to perceived threats including the emergence of collective action. But what shapes people's perception of the government and its actions?

People perceive today's societies as dominated by extreme and polarised views (Levendusky & Malhotra, 2016) combined with a lack of shared values (Ahler, 2014). This is not surprising given that news coverage—which is a key information source for citizens (Shehata & Strömbäck, 2022; Tung Lai et al., 2015)—frequently focuses on the ever-increasing polarisation of societies. Moreover, the news is dominated by the threats the world is facing, like wars, climate crises, and economic recessions, and governments' reactions to threats. Thus, people are constantly exposed to information about threats, societal polarisation, and government actions. The overarching question of the present article is how perceived threat and information about societal polarisation shape people's intention for collective action against the government. We explore whether perceptions of polarisation and perceived threat in combination increase (extreme) non-normative collective action intentions. We further examine trust in the government as a mediator. We test whether perceived threat and polarisation jointly lower political trust which, in turn, could enhance willingness to engage in (extreme) non-normative collective action.

1.1. Collective Action

Collective action describes any kind of activity that groups and individuals engage in to achieve a common goal. This happens usually in response to perceived injustice and aims to promote social change, to protect or enhance the status of the ingroup (Becker, 2012; Wright et al., 1990). Collective action can take many different forms, including, but not limited to, strikes, protests, and petitions. These actions include legal as well as illegal behaviour and can be categorised as normative collective action, non-normative collective action, or extreme non-normative collective action (Tausch et al., 2011). In most Western countries, non-violent forms of action are considered normative; however, social norms and context further influence what constitutes normative behaviour (Louis et al., 2005; Tausch et al., 2011). Non-normative actions contrast with societal norms like damage of personal property or disturbing society (Schuetze, 2023). Extreme non-normative collective action usually includes an element of violence directed against others (e.g., suicide bombings; Tausch et al., 2011). Additionally, being part of a radical or extremist group can be seen as a non-normative or even extreme non-normative action. Research on collective action also emphasises the role of antecedent factors such as social norms for collective action (van Zomeren et al.,

2008). Specifically, there exists now a growing body of empirical work that focuses on the psychological variables that precede extreme non-normative action (for a review see Feddes et al., 2023). For example, work by Doosje et al. (2012, 2013) showed that intergroup threat perceptions and perceived illegitimacy of authorities were related to more support for extreme, non-normative actions.

Collective action has a close but complex connection to political trust. Studies found that political trust increases political participation including normative collective action (Bäck & Christensen, 2016; Gulevich & Osin, 2023). In contrast, other studies found that trust is associated with less collective action (Bekmagambetov et al., 2018; Corcoran et al., 2011). Further, findings showed that lower trust was associated with increased likeliness to engage in non-normative action (Hooghe & Marien, 2013; Šerek & Macek, 2014), while higher trust was associated with higher compliance to governmental demands (Levi & Stoker, 2000). Combined, these findings demonstrate that collective action is closely linked to political trust, but the relationship depends on contextual factors. Here we focus on two contextual factors—perceived threat and polarisation—that might shape citizens' level of trust and in turn their willingness to engage in collective action.

1.2. Perceived Threat and Collective Action

Threats refer to the perception that any real or possible change in personal circumstances or social standing of one's group might result in negative consequences (Fritsche et al., 2011). They are associated with a wide range of social attitudes and behaviour like explicit and implicit prejudice (Aberson & Gaffney, 2009), right-wing mindset (Jost et al., 2017; Onraet et al., 2013, 2014), and identification with the ingroup (Wilson & Hugenberg, 2010). Threat perceptions play a crucial role in political behaviour, like citizens' voting behaviour and support for political policies (Brandt & Bakker, 2022). For example, experiencing relative group deprivation was associated with support for Donald Trump in the 2016 US election (Major et al., 2018), for Marine Le Pen in the 2012 French election (Urbanska & Guimond, 2018), and for Brexit (Macdougall et al., 2020). Holman et al. (2016) found that perceived terrorist threat reduced the support of a Democratic female leader in the US and a conservative female leader in the UK. People who experience personal economic hardship are less trustful of the government and less willing to adhere to government guidelines (Lobera et al., 2024).

Further, higher levels of perceived symbolic and realistic threat were associated with higher collective action intentions (Ayanian & Tausch, 2016; Ayanian et al., 2020), for example in the conflict between Turks and Kurds (Çakal et al., 2016) and pro-environmental behaviour in the face of climate threat (Stollberg & Jonas, 2021). In the context of Hong Kong, perceiving symbolic threat increased support and willingness to engage in normative and violent collective action (Chan et al., 2023; Cheung et al., 2022; Gulliver et al., 2023). More generally, threat perceptions increased pro-group behaviour and collective thinking (Çakal et al., 2016; Fritsche et al., 2017; Yustisia et al., 2020), which in turn might increase willingness to engage in collective action. In sum, these findings highlight the role of threats in political behaviour, which are associated with normative and non-normative collective action as a means to mitigate threats.

1.3. Perceived Polarisation and Collective Action

Threat perceptions do not take place in isolation but rather they interact with other factors related to the societal climate. A united and well-functioning society might be well-equipped to handle threats adequately, while a divided society might struggle to ward off threats. Thus, threat perceptions might differ depending on the levels of polarisation within societies and they might reinforce each other. For example, realistic threats like resource competition are associated with increased prejudice and outgroup hostility (Stephan & Stephan, 2000). This might lead to increased polarisation in society regarding issues like welfare benefits and migration policy. Indeed, events like financial, environmental, and refugee crises, which all can pose threats to the status quo, co-occur with rising levels of polarisation in Western societies.

Polarisation research distinguishes between different types of polarisation like actual versus perceived polarisation or affective versus issue polarisation. According to Lees and Cikara (2021), actual polarisation refers to the objective distance between individuals or groups, while perceived polarisation captures the subjectively perceived distance between individuals or groups. Thus, it focuses on the estimated difference between groups and captures a subjective experience of polarisation (Ahler, 2014; Keltner & Robinson, 1993; Lees & Cikara, 2021; Levendusky & Malhotra, 2016; Westfall et al., 2015). Further, affective polarisation captures the emotional component of polarisation. (i.e., how individuals feel about others in society), while issue polarisation captures the cognitive component of polarisation (i.e., how people feel about the issue). All these types of polarisation have been linked to political behaviour. Polarised citizens report more distrust and less political participation than non-polarised citizens (Fiorina & Abrams, 2008; Layman et al., 2006). Polarisation has been predominantly associated with negative outcomes like lower voter turnout, increased outgroup hostility, lower trust, and lower cooperative behaviour (Enders & Armaly, 2019; Lee, 2022). Yet, contrasting these assumptions, Smith et al. (2024) argued that polarisation can foster political engagement "which is itself critical for pluralistic societies in general, and democracies in particular" (p. 1). They postulated that polarisation could bring like-minded people together to challenge existing inequalities and advocate for positive social changes. As such, polarisation processes can be the underlying mechanism of collective action. The emerging form of collective action (normative versus non-normative) depends on the social contexts and its prevailing norms.

Most polarisation studies rely on measurements of polarisation, while the manipulation of polarisation has been less studied. Crimston et al. (2022) manipulated polarisation using a 60-second video about a fictional society which either highlighted consensus (low polarisation) or dissent (high polarisation) on societal issues. Across their studies they found that high polarisation was linked to perceptions of anomie and support of strong leaders. Anomie describes the perceived breakdown of social norms and rules, which might indicate that in a polarised society, people perceive norms as less valid. Similarly, Lee (2022) manipulated polarisation using fake newspaper articles that included information about the level of polarisation within society. They found that, in the low polarisation condition, cooperative behaviour and social trust were higher. There was no difference between the high polarisation and control condition without polarisation information. In combination, the findings indicate that information about polarisation on a societal level impacts individuals' behaviour and attitudes. We build on these findings by implementing a polarisation manipulation to explore its impact on trust in governments and (non-)normative collective action intentions.

1.4. The Current Research

The current research uses an experimental approach to test the joint effect of polarisation and threat on collective action intentions. We examine whether willingness to engage in (extreme) non-normative collective action is higher when societies are framed as being polarised. Thus, we manipulate the information participants receive about the level of societal polarisation on an issue. We further test how this relationship is impacted by repeated exposure to realistic threats. In this study, we manipulate the level of perceived realistic threat, i.e., perceived threats to a group's material well-being, such as threats to safety, resources, or jobs (Stephan & Stephan, 2000). In addition, the threat originates from the governing institution, resulting in a combination of realistic threat and political threat, i.e., the feeling that one's political power is undermined. We test three main hypotheses. First, building on the findings that in polarised environments norms are perceived as less valid and outgroup hostility is higher, we assume that (extreme) non-normative collective action intentions are higher in polarised climates:

> H1a: In a polarised society, inaction and normative collective action intentions are lower than in a non-polarised society.

> H1b: In a polarised society, non-normative collective action intentions are higher than in a non-polarised society.

Second, polarisation has been linked to lower trust, which in turn can increase collective action intentions, especially non-normative intentions. Thus, we assume that political trust mediates the relationship between polarisation and collective action:

> H2: In a polarised society, compared to a non-polarised society, trust in government is reduced, which in turn increases non-normative collective action intentions.

Lastly, we expect that perceived threat moderates the relationship between polarisation and trust. Realistic threat is associated with reduced trust in government. We assume that this link is even stronger in a polarised society compared to a non-polarised society:

> H3: The association between polarisation and trust in the government is stronger under threat.

2. Methodology

2.1. Pilot Study

To validate and choose the vignettes for the main study, we piloted 16 vignettes. We used a 2×2 repeated-measure design. The sample consisted of 48 student participants who received course credits for their participation. We excluded two participants who did not pass the attention check and four participants who spent significantly longer than average on the survey. The final sample consisted of 42 students (89.6% female; M_{age} = 20). After giving informed consent, each participant was randomly assigned to a threat (no threat vs threat) and a polarisation (no polarisation vs polarisation) condition. Each participant saw eight vignettes in randomised order and rated them on perceived level of threat and polarisation. The results

indicated that participants in the threat condition ($M = 5.79$, $SD = 0.54$) perceived significantly more threat than individuals in the no threat condition ($M = 2.97$, $SD = 1.19$; $t = -9.38$, $p < .001$). Participants in the polarisation condition ($M = 3.38$, $SD = 1.6$) perceived significantly more polarisation than individuals in the no polarisation condition ($M = 5.95$, $SD = 0.8$; $t = -6.59$, $p < .001$). There was no significant difference in perceived polarisation between the threat conditions ($M_{\text{no threat}} = 4.4$, $SD_{\text{no threat}} = 1.64$; $M_{\text{threat}} = 4.91$, $SD_{\text{threat}} = 1.95$; $t = -.912$, $p = .37$). There was no significant difference in perceived threat between the polarisation conditions ($M_{\text{no polarisation}} = 4.56$, $SD_{\text{no polarisation}} = 1.61$; $M_{\text{polarisation}} = 4.3$, $SD_{\text{polarisation}} = 1.82$; $t = 0.504$, $p = .617$). Based on the pilot we chose three vignettes for the main study. The three chosen vignettes capture realistic threat related to economic hardship (i.e., higher energy, lack of governmental support, and investment in the region), and showed on average a three-point difference between the threat and no threat condition. Detailed information about the pilot is displayed in the Supplementary File (Table S1 and S2).

2.2. Main Study

2.2.1. Sample and Design

We collected 1,004 UK participants via Prolific. As preregistered, we excluded one participant who did not consent, three participants who failed the attention checks, 15 participants who completed the study in less than five minutes, and three participants who took more than 60 minutes to complete the study. The average completion time was 13.29 minutes. In total, 21 participants were excluded, resulting in a final sample of 982 UK participants (485 women, 488 men, 8 non-binary, 1 other; $M_{\text{age}} = 44.58$, $SD_{\text{age}} = 14.62$). 43.08% ($N = 423$) did not have a university degree and 56.92% ($N = 559$) had an undergraduate degree or higher. We implemented a 2×2 repeated-measure design. Each participant was randomly allocated to a threat condition (no threat: $N = 482$ vs threat: $N = 491$) and a polarisation condition (no polarisation: $N = 486$ vs polarisation: $N = 487$).

After giving informed consent, participants were introduced to the fictitious country Bovenland (Feddes et al., 2024). They were informed that they would become citizens of Bovenland, either as Southerners or Northerners, and that Southerners have lower job opportunities, lower income, and less access to healthcare compared to Northerners. Using a mock minimal group task, all participants were assigned as Southerners. A manipulation check ensured that participants were able to remember which group they were assigned to (which was always the "Southerners"). To further engage them with the scenario, participants read Bovenland's constitution and indicated which laws they deemed most important. Then, participants were introduced to the vignette tasks with a practice task to familiarise themselves with the answering format. They were presented with three vignettes containing newspaper articles and indicated their response after each vignette. After completing the vignette tasks, participants filled out multiple questionnaires about their social and political attitudes, followed by demographics and full disclosure. All materials are displayed in the Supplementary Files.

2.2.2. Instruments

To manipulate perception of threat and polarisation, participants read newspaper articles in the *Bovenland Daily* about a situation that impacts them as a Southerner and the government's reaction to it. Depending on

the condition, the article included a realistic threat to their ingroup (to Southerners) or not. All participants saw the vignettes in the same order, starting with the least threatening to the most threatening according to the pilot study. The first vignette reported a new energy bill, which was rejected (no threat) or passed (threat). The second vignette reported a referendum related to governmental infrastructure investment. The referendum was either ignored, resulting in no governmental investment (threat), or it was taken into account, resulting in governmental investment (no threat). The third vignette reported a bill about releasing emergency funds, which was either passed (no threat) or rejected (threat). Additionally, the news article contained polls of the public opinion regarding the issue discussed. These polls either showed a divided society with extreme opinions on both ends of the distribution (polarised condition) or a society with normally distributed opinions (no polarisation condition).

To assess *collective action intentions*, participants responded to each of the three news scenarios by distributing 100 points across four different action options, ranging from inaction to extreme non-normative action. For example, the first threatening scenario negatively affected Southerners due to a new policy regarding energy costs. Using their 100 points, participants responded how likely they would be to either (a) do nothing and move on with their day, (b) sign a petition against the new policy, (c) block the entrance of a big energy producer, or (d) call for violently demonstrating in front of the members of parliament private homes. These answer options reflect the collective action strategies: (a) inaction, (b) normative action, (c) non-normative action, and (d) extreme non-normative action, respectively. The wording of the items was based on previous work by Tausch et al. (2011) where the different forms of collective action were described and measured. For the present purpose, we adapted the questions about collective action intentions to the specific context. The allocated points to each option were used as dependent variables ranging from 0 to 100 possible allocated points.

Trust in the government and the people in Bovenland was measured after the last vignette using two items adapted from the European Social Survey (Curtice & Bryson, 2001). Participants responded on a scale from 0 (*you can't be too careful*) to 10 (*most people can be trusted*). The items read: "As a Southerner living in Bovenland, would you say that most people can be trusted, or that you can't be too careful in dealing with people?" and "As a Southerner living in Bovenland, to what extent would you trust the government of Bovenland?" The items were highly correlated ($r = .51$, $p < .001$).

Participants indicated their *political orientation* on a 11-point scale from −5 (*progressive*) to 5 (*conservative*) regarding economic and cultural issues separately (Choma et al., 2009). The items were highly correlated ($r = .65$, $p < .001$) and combined into one score ($M = -0.87$, $SD = 2.23$).

At the end of the survey, participants reported their gender, age, and education level.

2.2.3. Data Availability

All confirmatory hypotheses and analyses for this project were pre-registered on AsPredicted; data and syntax are available online (https://osf.io/q9n5z/?view_only=8dbc5411f2c34ca1abe48919f6feae39)

2.2.4. Sample Size

A priori sample size calculation using GPower for a repeated measurement design with a between factor interaction indicated that 952 participants are required to find a small interaction effect ($f = 0.1$) with a power of .90, alpha = .05 and correlation of $r = .5$ between repeated measures. To account for individuals who fail attention checks, we aimed to collect 1000 participants.

3. Results

3.1. Descriptives

To obtain a first impression of our data, we explored the means of the different collective action intentions and trust per time point (Table 1). The mean points assigned to extreme non-normative action intention were very low. Extreme non-normative action intentions and non-normative action intention correlated positively ($r = .21$, $p < .001$) and thus we combined them into one composite score. Political trust was positively associated with inaction ($r = .21$, $p < .001$) and negatively with all forms of collective action intentions, especially strong with non-normative action intentions ($r = -.26$, $p < .001$). Societal trust showed the same pattern, but the correlations were weaker and, in the case of normative action, non-significant. Women and older participants were more likely to indicate normative action intentions and men and younger participants more likely to indicate non-normative action intentions (respectively $r = .16$, $p < .001$ and $r = .08$, $p < .05$).

Table 1. Descriptive.

Variable	mean	sd	median	min	max	skew	kurtosis	se
Inaction Mean	39.41	26.41	33.67	0	100	0.40	−0.63	0.84
Inaction M1	36.37	36.89	22.00	0	100	0.61	−1.14	1.18
Inaction M2	62.98	37.54	75.00	0	100	−0.48	−1.32	1.20
Inaction M3	18.87	27.69	5.00	0	100	1.62	1.72	0.88
Normative Action Mean	44.32	21.26	44.50	0	100	−0.07	−0.40	0.68
Normative Action M1	54.47	34.81	59.00	0	100	−0.22	−1.26	1.11
Normative Action M2	26.36	29.97	16.00	0	100	0.92	−0.27	0.96
Normative Action M3	52.13	29.22	50.00	0	100	−0.06	−0.80	0.93
Non-Normative Action Mean	12.78	13.37	9.50	0	66.67	1.17	1.04	0.43
Non-Normative Action M1	5.88	13.01	0.00	0	100	2.76	8.35	0.42
Non-Normative Action M2	6.50	14.63	0.00	0	100	3.15	11.92	0.47
Non-Normative Action M3	25.97	25.91	20.00	0	100	0.87	0.09	0.83
Extreme Non-Normative Action Mean	3.49	8.43	0.00	0	74.67	3.71	17.97	0.27
Extreme Non-Normative Action M1	3.29	9.78	0.00	0	77	3.87	16.34	0.31
Extreme Non-Normative Action M2	4.16	12.68	0.00	0	100	4.36	22.68	0.40
Extreme Non-Normative Action M3	3.03	10.99	0.00	0	100	5.34	34.10	0.35
Political Trust	2.97	2.43	2.00	0	10	0.69	−0.50	0.08
Societal Trust	4.63	2.45	5.00	0	10	0.00	−0.85	0.08

3.2. Hypotheses 1a and 1b

3.2.1. Preregistered Repeated Measurement ANOVA

In the experiment, participants' collective action intentions were measured three times, i.e., after each vignette. Therefore, we preregistered a two-factorial repeated-measure ANOVA to explore the effect on repeated exposure to threat and polarisation on collective action intentions. Focusing on inaction as the dependent variable, we found a significant main effect for threat ($F(1,969) = 65.90$, $p < .001$, generalised $\eta^2 = .04$) and measurement point ($F(2,1938) = 697.52$, $p < .001$, generalised $\eta^2 = .23$) as well as significant interactions between polarisation and measurement point ($F(2,1938) = 8.28$, $p < .001$, generalised $\eta^2 = .003$), threat and measurement point ($F(2,1938) = 25.63$, $p < .001$, generalised $\eta^2 = .02$), and a significant three-way interaction ($F(2,1938) = 4.13$, $p = .015$, generalised $\eta^2 = .001$). The difference between points allocated to inaction between measurement was larger in the no threat condition compared to the threat condition. The same pattern emerged for polarisation; there were larger differences in the no polarisation compared to the polarisation condition. The effect of measurement point is not linear, that is the likelihood of inaction does not show a linear change. Rather it seems like interactions with measurement point were driven by Vignette 2. The results for normative and non-normative action intentions showed a similar pattern. Detailed description in the Supplementary Files. In sum, the results indicated that Vignette 2 had a different effect on collective action intentions than Vignette 1 and Vignette 3. Therefore, in the following analyses, we decided to average the collective action intentions across the three measurement points. We discuss the implications of the different responses to the vignettes in the general discussion.

3.2.2. Two-Way Factorial ANOVA

To further test our first hypotheses (H1a and H1b), we conducted three separate two-way factorial ANOVAs with inaction, normative, and non-normative collective action intentions (including extreme non-normative collective action intentions) averaged across the two measurement points as dependent variables while controlling for gender, age, and political orientation. Detailed results are in the Supplementary Files.

For inaction as the dependent variable, we found a significant main effect of threat ($F(1,962) = 34.90$, $p > .001$, partial $\eta^2 = .03$), but no main effect of polarisation and no interaction effect between threat and polarisation. The results indicated that under threat ($M = 32.76$, $SD = 25.89$) compared to no threat ($M = 46.19$, $SD = 25.21$), participants moved away from inaction to any other form of action.

For normative collective action intentions, we found a significant main effect of threat ($F(1,962) = 10.86$, $p = .001$, partial $\eta^2 = .01$) and a significant main effect of polarisation ($F(1,962) = 4.31$, $p = .038$, partial $\eta^2 = .00$). We found no interaction effect between threat and polarisation. The results showed that under threat ($M = 47.29$, $SD = 20.83$) or if society is perceived to be polarised ($M = 46.01$, $SD = 21.34$), participants were more likely to engage in normative action compared to no threat ($M = 41.29$, $SD = 21.29$) or in a cohesive society ($M = 42.62$, $SD = 21.07$).

We found a main effect of threat ($F(1,962) = 24.21$, $p < .001$, partial $\eta^2 = .02$), but no main effect of polarisation and no interaction effect. In the threat condition, participants reported significantly more intention of non-normative and extreme non-normative action ($M = 14.07$, $SD = 13.49$) compared to the no threat condition ($M = 11.51$, $SD = 13.13$).

Taken together, it seems that threat primarily drives people to move toward collective action and also to (extreme) non-normative forms of collective action. Polarisation did not have a consistent effect on collective action intentions. Based on these results, we reject H1a and H1b. We did not find a main effect of polarisation for inaction and non-normative action intentions, and for normative action intentions we found the reversed pattern. Normative action intentions were higher in the polarised climate than in the non-polarised climate. Further, the results indicate that threat is a driver for normative and non-normative behaviour independent of the level of polarisation.

3.3. Hypotheses 2: Mediation Analyses

Next, to explore whether polarisation impacts collective action intentions indirectly through reducing trust in governments, we conducted mediation analyses with trust in the government as a mediator between polarisation and inaction (model 1), normative (model 2) and non-normative collective action intentions (model 3; see Figure 1). We controlled for gender, age, and political orientation. Detailed results are in the Supplementary Files. We found that polarisation increased trust in governments. For inaction (model 1) there was no significant indirect effect of polarisation on inaction via trust in government. For normative action (model 2) we found a significant indirect effect of polarisation on normative action intentions partially mediated through trust in the government. It was a suppression effect, indicating that after controlling for trust, the effect of polarisation on normative action was stronger than without controlling for trust. For non-normative action intentions (model 3), we found a significant indirect effect of polarisation on non-normative action intentions partially mediated through trust in the government. In a polarised climate, trust in government increases, which in turn reduces non-normative action intentions. Based on the results, we reject H2. While political trust mediates the link between polarisation and collective action intentions, the link between polarisation and trust was opposed to our assumption. Polarisation increased political trust, instead of decreasing it. We additionally explored societal trust as a mediator. Polarisation was not associated with societal trust and thus there was no indirect effect on collective action intentions mediated through societal trust. Potential reasons are discussed in the general discussion section.

3.4. Hypotheses 3: Moderated Mediation Analyses

Next, we included threat and the interaction between polarisation and threat in the models (models 4–6, Figures 2–4). We controlled for gender, age, and political orientation. Detailed results are in the Supplementary Files. For inaction (model 4), we found a direct effect of polarisation and threat. There was a significant effect of threat on trust in government: Trust in the government was higher in the no threat condition compared to the threat condition. There was also a significant direct effect of polarisation on trust. That is, in the polarisation condition, trust in government was higher than in the no polarisation condition. There was a significant interaction between polarisation and threat on trust. Simple effects indicated that polarisation only impacted trust in the no-threat condition. This also impacted the indirect effect. That is, in the no-threat condition, polarisation increased trust in government which in turn reduced collective action intentions.

For normative action intentions (model 5), we found a direct effect of polarisation and threat on normative action, but no significant interaction between polarisation and threat on collective action intentions. For both threat and polarisation there was an indirect effect on normative action intentions, but in opposing directions.

Model 1
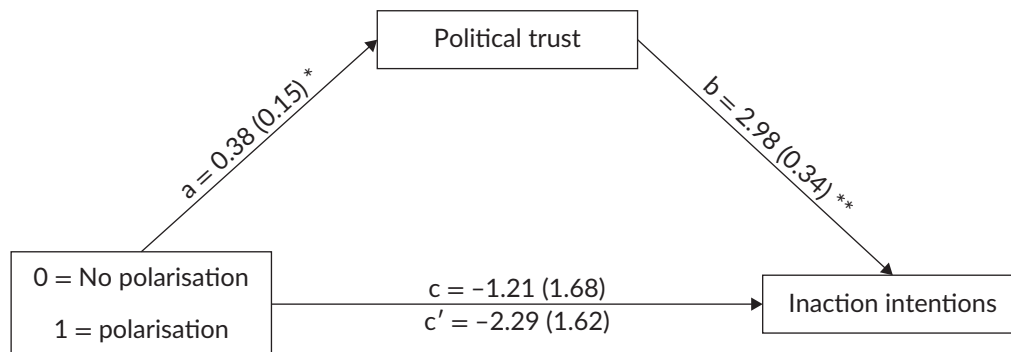
Model 2
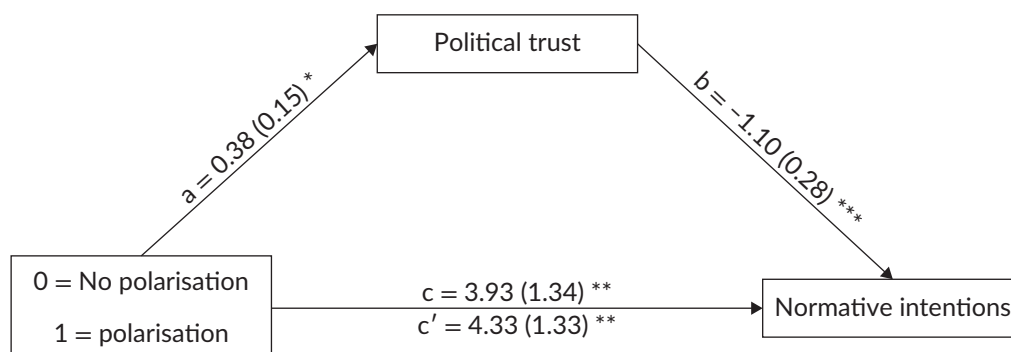
Model 3
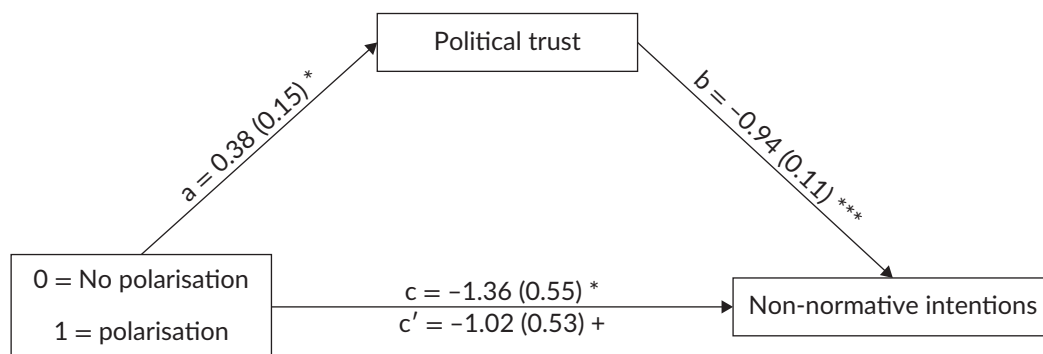
Figure 1. Mediation model for inaction, normative action, and non-normative action.

That is, for threat, the link with normative action intentions got weaker after including trust, but for polarisation the link got stronger. For non-normative action intentions (model 6), we found a direct effect of threat, but no direct effect of polarisation and no significant interaction between polarisation and threat. There was an indirect effect of threat on non-normative action intentions. That is, in the threat condition trust in government decreased which in turn increased non-normative action intentions.

In a nutshell, the results suggest a positive indirect effect: Polarisation increases trust in government, and this is related to more inaction, and less normative action intentions, but only in the low-threat condition. In the high-threat condition, there was no significant indirect effect.
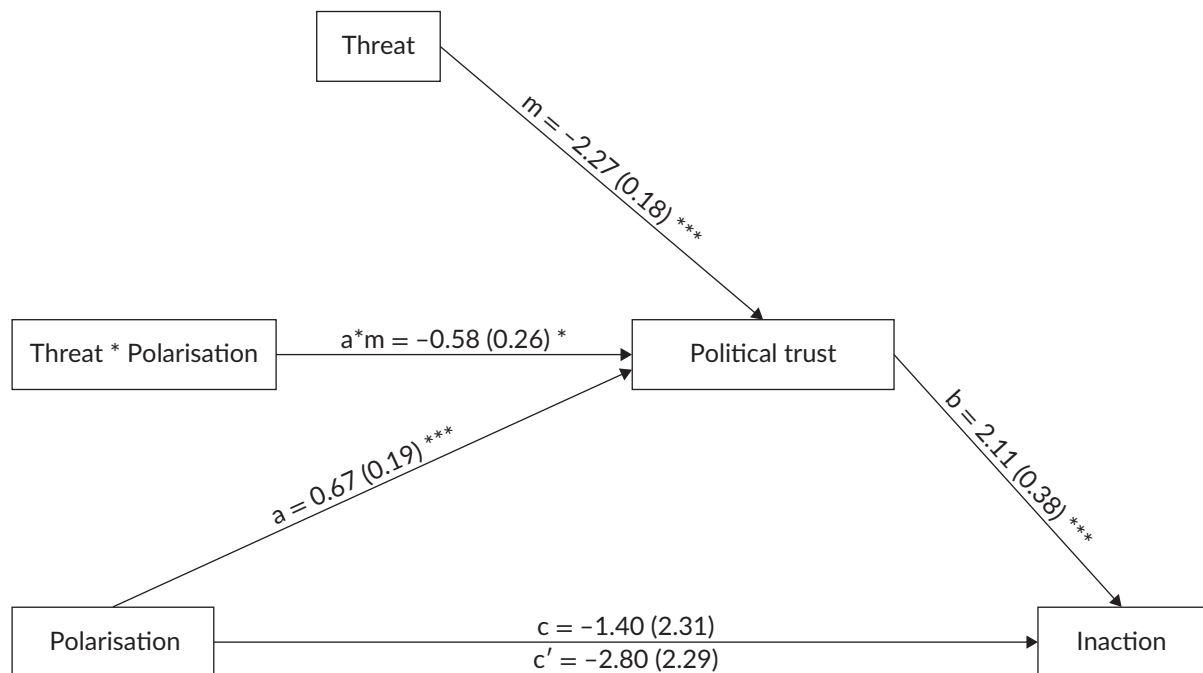
Figure 2. Moderated mediation for inaction as dependent variable (model 4).
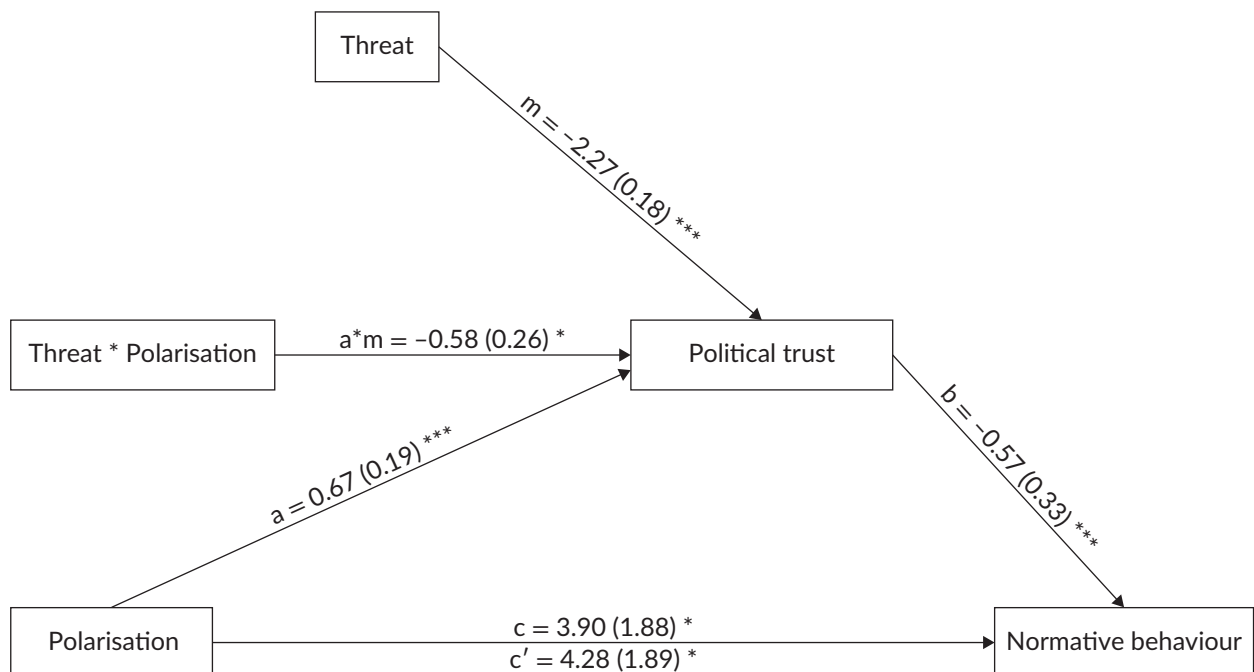
Figure 3. Moderated mediation for normative action as dependent variable (model 5).
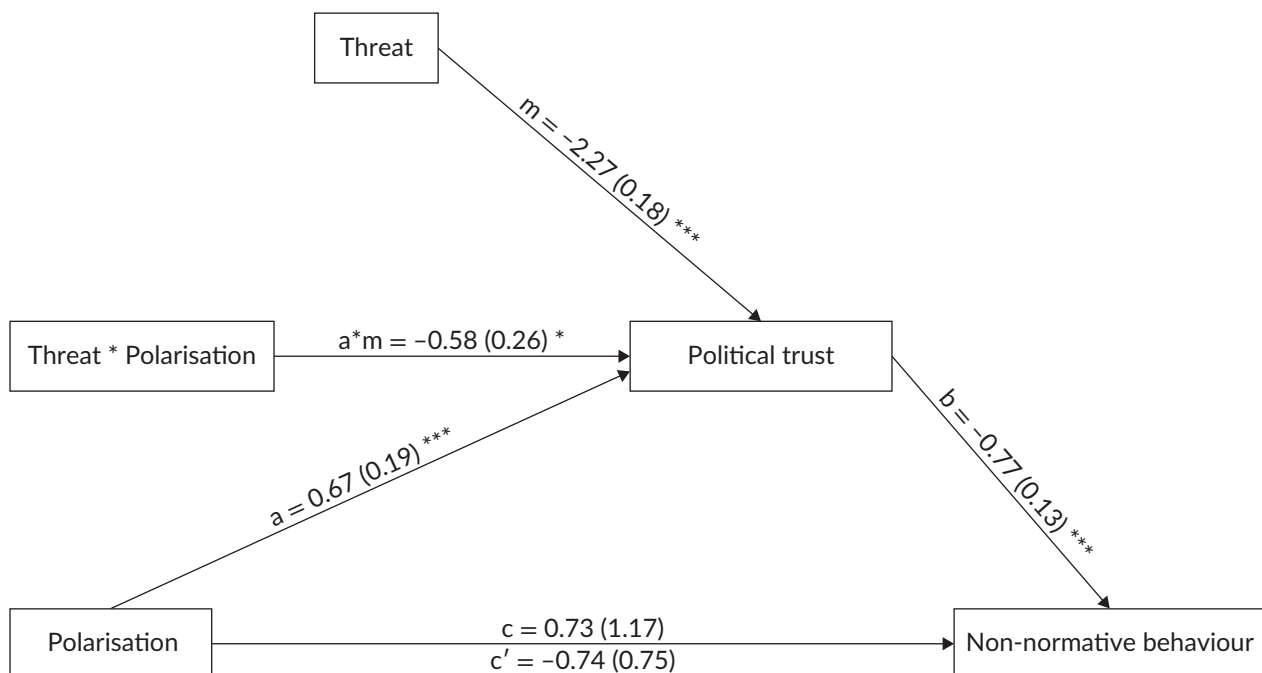
Figure 4. Moderated mediation for non-normative action as dependent variable (model 6).

4. General Discussion

Collective action emerges as a reaction to perceived threats including perceived disadvantaged based on one's own group membership. The nature of collective action (e.g., normative or non-normative behaviour) depends on the societal climate. It serves as a benchmark to evaluate what is acceptable and desirable within a society (Chiu et al., 2010). This article explores two factors that could impact the perception of the societal climate: perceived polarisation (i.e., the extent to which society is polarised about an issue) and perceived threat.

We found that collective action intentions were predominately driven by threat perceptions. Under threat (here realistic threat to the ingroup) participants engaged in more collective action in general, including (extreme) non-normative action. This aligns with research exploring the impact of economic inequality at a societal level on people's behaviour. When economic inequality at a societal level is low, people tend to be more cooperative and benevolent. When inequality is high, people act more aggressively and individualistically towards those who are better off (Sánchez-Rodríguez et al., 2019). Sanchez-Rodrigues and colleagues proposed that a shift in social norms postulates that what is acceptable in an *unequal* context is not acceptable in an *equal* context. In line with this, we found that under conditions of threat to the ingroup, (extreme) non-normative collective action responses became more acceptable to our participants.

Polarisation perceptions, however, did not consistently impact collective action. Instead, our findings indicate that societal polarisation is most impactful when no threat is present. In the no-threat condition, and in contrast to our hypotheses, polarisation even increased trust in government and reduced normative collective action intentions. These findings were surprising and counterintuitive. However, multiple plausible explanations exist.

First, unlike previous studies, we did not measure personal polarisation, but rather manipulated polarisation in the society. Therefore, it is conceivable that the perceived polarisation of the society acts differently than

individualised polarisation. While the latter might motivate collective action, a seemingly polarised society might lead people to shift their trust towards the government. As perceptions of polarisation increase and individuals' trust in others decreases, trust in the government might increase. In other words, polarisation might lead individuals to seek support beyond their peer-citizens, leading to increased trust in governing institutions. We find, however, that this only holds if the government is perceived to be a potential alternative to seek support from. In the threat condition, the government acts in opposition to the participants and this compensatory trust response to polarisation is not shown.

Secondly, the effect of threat in our experiment could have superseded any effects of the polarisation condition. In our study, we focused on one type of threat: realistic threat. Since the fabricated threats focused on unequal treatments resulting in economic hardship for one group, inducing threat might inevitably lead to increasing polarisation, even if not explicitly stated. That is, if two groups are competing against each other, one might automatically believe that society is polarised. We did not see an impact of threat manipulation on polarisation perception in our pilot study, which speaks against the notion that realistic threat automatically triggers polarisation. Nonetheless, in future studies, it would be interesting to see if external threats unrelated to intergroup relations, such as environmental catastrophes or internal threats like loss of control, would result in a different pattern.

Third, polarisation was manipulated by displaying societal attitudes using different distributions. We relied on people understanding the graphs, and while the pilot indicated that the polarisation manipulation worked, there is research highlighting individual differences in understanding graphs and interpreting them correctly (Garcia-Retamero & Galesic, 2010). Unlike previous work (Lee, 2022), we did not mention polarisation explicitly. Rather, the distribution was just part of the newspaper. As such, it might have been less focal to people's perception of the article, or its impact might be overshadowed by the threat manipulation. Future research could test different methods of presenting information about polarisation.

We exposed participants to three vignettes exploring whether collective action intentions increase over time. However, our results indicate that there was no linear increase, instead Vignette 2 seemed to trigger a different reaction. A potential reason for this is a difference in how threat was formulated in the different vignettes. In Vignette 1 and Vignette 3, the government introduced a bill that would disproportionally impact their group in the future. As such, participants might have opted for collective action to reverse the newly introduced bill and maintain the status quo. In Vignette 2 their group was already disadvantaged and the government rejected a bill that would reduce the disadvantage. Therefore, the starting points in the vignettes were different: an existing equality that is dismantled compared to an inequality that is maintained. A potential explanation could be system-justifying beliefs, which can result in disadvantaged groups supporting, or at least not challenging, unfair systems (Jost & Hunyady, 2003). Additionally, in Vignette 2, a referendum was explicitly ignored by the government. While all vignettes scored high on perceived threat in the pilot, the type of perceived threat might differ. Future studies are needed to understand the role of threat types and status quo (e.g., introducing vs maintaining an inequality) for collective action intentions.

Our results highlight the crucial role political trust plays in collective action intentions. If trust in government is low, non-normative and extreme collective action intention may be more likely to occur. Indeed, this is in line with previous work by Doosje et al. (2012, 2013) who found in a cross-sectional survey study that greater perceived legitimacy of the authorities, which included the government, is related to more support of

ideology-based violence by others, and stronger intentions to use ideology-based violence toward outgroup members. Importantly, in the present study, we did not find evidence that greater polarisation perceptions are related to greater distrust in authorities. In that sense, the present social context as introduced in the paradigm seems not to have created greater affective polarisation (Iyengar & Krupenkin, 2018; Waldrop, 2021). Future studies could examine whether a social context where there are clear divisions between outgroups based on distrust would affect extreme non-normative collective action aimed at an outgroup.

One limitation of the current study is the use of a fictional society. Future studies could test the external validity by using, for example, experience sampling methods to explore how frequently people experience threats and polarisation in everyday life and how it relates to collective action behaviour. Moreover, this approach would allow testing the possible reinforcing relationship between polarisation, threat, and forms of collective action over time. For example, the visibility of non-normative collective actions might influence perceived polarisation which in turn might increase threat perception. Higher threat perception might result in more non-normative action, triggering a vicious circle that could destabilise society and democratic governance.

Besides the question of external validity, there is also the question of generalisability to the UK and beyond. We do not know which part of the UK people are from. It is probably predominately England, and potentially UK citizens from Scotland, Wales or Northern Ireland would show a different pattern. Generalising the findings to other countries including Scotland, Wales, and Northern Ireland, might be problematic. On the one hand, the paradigm is fictional and as such does not rely on a specific cultural aspect. On the other hand, the vignettes, behaviours, and underlying norms do refer to a Western cultural context. As mentioned in the introduction, what constitutes normative behaviour depends on the context. The same can be said for threat and polarisation perceptions. People from Northern Ireland, for example, with its history, might react differently to perceived threats and polarisation resulting than people from England. Therefore, whether the findings can be generalised to other contexts remains an open question to be explored in the future.

5. Conclusion

The present study suggests that under threat people are mobilised, resulting in higher normative and (extreme) non-normative collective action tendencies to improve their situation. Levels of polarisation within the society, however, do not consistently relate to collective action intentions. Rather, it seems that polarisation has a more nuanced impact: It matters in low-threat environments, but not when threat is high.

Funding

Conflict of Interests
The authors declare no conflict of interests.

Data Availability
Data and syntax are available at: https://osf.io/q9n5z/?view_only=8dbc5411f2c34ca1abe48919f6feae39

Supplementary Material

Supplementary material for this article is available online in the format provided by the author (unedited).

References

110 police officers injured in UK riots so far as real Brit patriots thank cops with sweet cards. (2024, August 7). *Mirror*.

Aberson, C. L., & Gaffney, A. M. (2009). An integrated threat model of explicit and implicit attitudes. *European Journal of Social Psychology*, *39*(5), 808–830. https://doi.org/10.1002/ejsp.582

Ahler, D. J. (2014). Self-fulfilling misperceptions of public polarization. *Journal of Politics*, *76*(3), 607–620. https://doi.org/10.1017/S0022381614000085

Ayanian, A. H., & Tausch, N. (2016). How risk perception shapes collective action intentions in repressive contexts: A study of Egyptian activists during the 2013 post-coup uprising. *British Journal of Social Psychology*, *55*(4). https://doi.org/10.1111/bjso.12164

Ayanian, A. H., Tausch, N., Acar, Y. G., Chayinska, M., Cheung, W. Y., & Lukyanova, Y. (2020). Resistance in repressive contexts: A comprehensive test of psychological predictors. *Journal of Personality and Social Psychology*, *120*(4), 912–939. https://doi.org/10.1037/pspi0000285

Bäck, M., & Christensen, H. S. (2016). When trust matters—A multilevel analysis of the effect of generalized trust on political participation in 25 European democracies. *Journal of Civil Society*, *12*(2). https://doi.org/10.1080/17448689.2016.1176730

Becker, J. C. (2012). The system-stabilizing role of identity management strategies: Social creativity can undermine collective action for social change. *Journal of Personality and Social Psychology*, *103*(4), 647–662. https://doi.org/10.1037/a0029240

Bekmagambetov, A., Wagner, K. M., Gainous, J., Sabitov, Z., Rodionov, A., & Gabdulina, B. (2018). Critical social media information flows: Political trust and protest behaviour among Kazakhstani college students. *Central Asian Survey*, *37*(4). https://doi.org/10.1080/02634937.2018.1479374

Boeren dumpen puin op wegen, stichten brand; alle wegen weer vrijgegeven. (2024a, February 6). *NOS Nieuws*. https://nos.nl/artikel/2507731-boeren-dumpen-puin-op-wegen-stichten-brand-alle-wegen-weer-vrijgegeven

Brandt, M. J., & Bakker, B. N. (2022). The complicated but solvable threat–politics relationship. *Trends in Cognitive Sciences*, *26*(5), 368–370. https://doi.org/10.1016/j.tics.2022.02.005

Breeden, A., & Specia, M. (2019, August 26). Dispute over Amazon gets personal for Bolsonaro and Macron. *The New York Times*.

Çakal, H., Hewstone, M., Güler, M., & Heath, A. (2016). Predicting support for collective action in the conflict between Turks and Kurds: Perceived threats as a mediator of intergroup contact and social identity. *Group Processes and Intergroup Relations*, *19*(6), 732–752. https://doi.org/10.1177/1368430216641303

Chan, C. S., Gulliver, R. E., Awale, A., Tam, K. Y. Y., & Louis, W. R. (2023). The influence of perceived threat and political mistrust on politicized identity and normative and violent nonnormative collective action. *Journal of Social and Political Psychology*, *11*(1), 126–144. https://doi.org/10.5964/jspp.7979

Cheung, H. K., Baranik, L. E., Burrows, D., & Ashburn-Nardo, L. (2022). Hiring discrimination against refugees: Examining the mediating role of symbolic and realistic threat. *Journal of Vocational Behavior*, *138*. https://doi.org/10.1016/j.jvb.2022.103765

Chiu, C.-Y., Gelfand, M. J., Yamagishi, T., Shteynberg, G., & Wan, C. (2010). Intersubjective culture. *Perspectives on Psychological Science*, *5*(4), 482–493. https://doi.org/10.1177/1745691610375562

Choma, B. L., Busseri, M. A., & Sadava, S. W. (2009). Liberal and conservative political ideologies: Different

routes to happiness? *Journal of Research in Personality*, *43*(3), 502–505. https://doi.org/10.1016/j.jrp.2008.12.016

Corcoran, K. E., Pettinicchio, D., & Young, J. T. N. (2011). The context of control: A cross-national investigation of the link between political institutions, efficacy, and collective action. *British Journal of Social Psychology*, *50*(4). https://doi.org/10.1111/j.2044-8309.2011.02076.x

Crimston, C. R., Selvanathan, H. P., & Jetten, J. (2022). Moral polarization predicts support for authoritarian and progressive strong leaders via the perceived breakdown of society. *Political Psychology*, *43*(4), 671–691. https://doi.org/10.1111/pops.12787

Curtice, J., & Bryson, C. (2001). *European social survey core questionnaire development—Chapter 6: The measurement of socio-political orientations*. European Social Survey. https://www.europeansocialsurvey.org/sites/default/files/2023-06/ESS_core_questionnaire_socio_political_ori.pdf

Doosje, B., Loseman, A., & van den Bos, K. (2013). Determinants of radicalization of Islamic youth in the Netherlands: Personal uncertainty, perceived injustice, and perceived group threat. *Journal of Social Issues*, *69*(3), 586–604. https://doi.org/10.1111/josi.12030

Doosje, B., van den Bos, K., Loseman, A., Feddes, A. R., & Mann, L. (2012). "My ingroup is superior!": Susceptibility for radical right-wing attitudes and behaviors in Dutch youth. *Negotiation and Conflict Management Research*, *5*(3), 253–268. https://doi.org/10.1111/j.1750-4716.2012.00099.x

Enders, A. M., & Armaly, M. T. (2019). The differential effects of actual and perceived polarization. *Political Behavior*, *41*(3), 815–839. https://doi.org/10.1007/s11109-018-9476-2

Feddes, A., Nickolson, L., van Bergen, N. R. J., Mann, L., & Doosje, B. (2023). Extremist thinking and doing: A Systematic literature study of empirical findings on factors associated with (de)radicalisation processes. *International Journal of Developmental Sciences*, *17*(1/3), 7–18. https://doi.org/10.3233/DEV-230345

Feddes, A., Ergin, A., Van Kurpershoek, R., Gligorić, V., & Rutjens, B. T. (2024). *From inaction to (extreme) action in Bovenland: An experimental study on the dynamics between social injustice, normative and (extreme) non-normative collective action*. Association of Social Psychological Researchers (ASPO).

Fiorina, M. P., & Abrams, S. J. (2008). Political polarization in the American public. *Annual Review of Political Science*, *11*, 563–588. https://doi.org/10.1146/annurev.polisci.11.053106.153836

Fritsche, I., Jonas, E., & Kessler, T. (2011). Collective reactions to threat: Implications for intergroup conflict and for solving societal crises. *Social Issues and Policy Review*, *5*(1), 101–136. https://doi.org/10.1111/j.1751-2409.2011.01027.x

Fritsche, I., Moya, M., Bukowski, M., Jugert, P., de Lemus, S., Decker, O., Valor-Segura, I., & Navarro-Carrillo, G. (2017). The Great Recession and group-based control: Converting personal helplessness into social class ingroup trust and collective action. *Journal of Social Issues*, *73*(1), 117–137. https://doi.org/10.1111/josi.12207

Garcia-Retamero, R., & Galesic, M. (2010). Who profits from visual aids: Overcoming challenges in people's understanding of risks. *Social Science and Medicine*, *70*(7). https://doi.org/10.1016/j.socscimed.2009.11.031

Gulevich, O. A., & Osin, E. N. (2023). Generalized trust and military attitudes in Russia: The role of national and global human identification. *British Journal of Social Psychology*, *62*(3). https://doi.org/10.1111/bjso.12648

Gulliver, R., Chan, C. S., Tam, K. Y. Y., Lau, I. S. K., Hong, Y. Y., & Louis, W. R. (2023). Political distrust, perceived threat, and intentions to engage in normative and violent collective action: A mixed-methods study. *European Journal of Social Psychology*, *53*(2), 401–417. https://doi.org/10.1002/ejsp.2910

Holman, M. R., Merolla, J. L., & Zechmeister, E. J. (2016). Terrorist threat, male stereotypes, and candidate evaluations. *Political Research Quarterly*, *69*(1), 134–147. https://doi.org/10.1177/1065912915624018

Hooghe, M., & Marien, S. (2013). A comparative analysis of the relation between political trust and forms of political participation in Europe. *European Societies*, *15*(1). https://doi.org/10.1080/14616696.2012.692807

Iyengar, S., & Krupenkin, M. (2018). The strengthening of partisan affect. *Political Psychology*, *39*, 201–218. https://doi.org/10.1111/pops.12487

Jetten, J., Mols, F., & Selvanathan, H. P. (2020). How economic inequality fuels the rise and persistence of the yellow vest movement. *International Review of Social Psychology*, *33*(1). https://doi.org/10.5334/irsp.356

Jost, J., Becker, J., Osborne, D., & Badaan, V. (2017). Missing in (collective) action: Ideology, system justification, and the motivational antecedents of two types of protest behavior. *Current Directions in Psychological Science*, *26*(2), 99–108. https://doi.org/10.1177/0963721417690633

Jost, J., & Hunyady, O. (2003). The psychology of system justification and the palliative function of ideology. *European Review of Social Psychology*, *13*(1), 111–153. https://doi.org/10.1080/10463280240000046

Keltner, D., & Robinson, R. J. (1993). Imagined ideological differences in conflict escalation and resolution. *The International Journal of Conflict Management*, *4*(3), 249–262. https://doi.org/10.1108/eb022728

Layman, G. C., Carsey, T. M., & Horowitz, J. M. (2006). Party polarization in American politics: Characteristics, causes, and consequences. *Annual Review of Political Science*, *9*. https://doi.org/10.1146/annurev.polisci.9.070204.105138

Lee, A. H. Y. (2022). Social trust in polarized times: How perceptions of political polarization affect Americans' trust in each other. *Political Behavior*, *44*(3), 1533–1554. https://doi.org/10.1007/s11109-022-09787-1

Lees, J., & Cikara, M. (2021). Understanding and combating misperceived polarization. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *376*(1822). https://doi.org/10.1098/rstb.2020.0143

Levendusky, M. S., & Malhotra, N. (2016). (MIS)perceptions of partisan polarization in the American public. *Public Opinion Quarterly*, *80*, 378–391. https://doi.org/10.1093/poq/nfv045

Levi, M., & Stoker, L. (2000). Political trust and trustworthiness. *Annual Review of Political Science*, *3*. https://doi.org/10.1146/annurev.polisci.3.1.475

Lobera, J., Santana, A., & Gross, C. (2024). Are we looking at crises through polarized lenses? Predicting public assessments of the official early responses to the COVID-19 pandemic in eight countries. *European Sociological Review*, *40*(1), 1–13. https://doi.org/10.1093/esr/jcad016

Louis, W. R., Taylor, D. M., & Douglas, R. L. (2005). Normative influence and rational conflict decisions: Group norms and cost-benefit analyses for intergroup behavior. *Group Processes and Intergroup Relations*, *8*(4). https://doi.org/10.1177/1368430205056465

Macdougall, A. I., Feddes, A. R., & Doosje, B. (2020). "They've put nothing in the pot!": Brexit and the key psychological motivations behind voting "remain" and "leave." *Political Psychology*, *41*(5), 979–995. https://doi.org/10.1111/pops.12660

Major, B., Blodorn, A., & Major Blascovich, G. (2018). The threat of increasing diversity: Why many White Americans support Trump in the 2016 presidential election. *Group Processes and Intergroup Relations*, *21*(6), 931–940. https://doi.org/10.1177/1368430216677304

Onraet, E., Dhont, K., & Van Hiel, A. (2014). Longitudinal relationship between threat and right-wing attitudes. The relationships between internal and external threat and right-wing attitudes: A three-wave longitudinal study. *Personality and Social Psychology Bulletin*, *40*, 712–725.

Onraet, E., Van Hiel, A., Dhont, K., & Pattyn, S. (2013). Internal and external threat in relationship with right-wing attitudes. *Journal of Personality*, *81*(3), 233–248. https://doi.org/10.1111/jopy.12011

Politie: niet-vreedzame demonstranten kaapten UvA-protesten. (2024b, November 25). *NOS Nieuws*.

Sánchez-Rodríguez, Á., Willis, G. B., Jetten, J., & Rodríguez-Bailón, R. (2019). Economic inequality enhances

inferences that the normative climate is individualistic and competitive. *European Journal of Social Psychology*, *49*(6), 1114–1127. https://doi.org/10.1002/ejsp.2557

Scenes from the student protests churning across the country. (2024, April 30). *The New York Times*.

Schuetze, C. F. (2023, January 2). Climate activists block German roads, raising awareness and animus. *The New York Times*.

Šerek, J., & Macek, P. (2014). Antecedents of political trust in adolescence: Cognitive abilities and perceptions of parents. *Journal of Applied Developmental Psychology*, *35*(4). https://doi.org/10.1016/j.appdev.2014.05.001

Shehata, A., & Strömbäck, J. (2022). Media use and societal perceptions: The dual role of media trust. *Media and Communication*, *10*(3), 146–157. https://doi.org/10.17645/MAC.V10I3.5449

Smith, L. G. E., Thomas, E. F., Bliuc, A.-M., & McGarty, C. (2024). Polarization is the psychological foundation of collective engagement. *Communications Psychology*, *2*(1). https://doi.org/10.1038/s44271-024-00089-2

Stephan, W. G., & Stephan, C. W. (2000). An integrated threat theory of prejudice. In S. Oskamp (Ed.), *Reducing Prejudice and Discrimination* (pp. 23–46). Lawrence Erlbaum Associated Publishers.

Stollberg, J., & Jonas, E. (2021). Existential threat as a challenge for individual and collective engagement: Climate change and the motivation to act. *Current Opinion in Psychology*, *42*. https://doi.org/10.1016/j.copsyc.2021.10.004

Tausch, N., Becker, J. C., Spears, R., Christ, O., Saab, R., Singh, P., & Siddiqui, R. N. (2011). Explaining radical group behaviour: Developing emotion and efficacy routes to normative and non-normative collective action. *Journal of Personality and Social Psychology*, *101*, 129–148. https://doi.org/10.1037/a0022728

Lai, W. T., Chung, W. C., & Sze, N. P. (2015). How do media shape perceptions of social reality? A review on cultivation theory. *Journal of Communication and Education*, *2*(2), 8–17. http://www.hkaect.org/jce/2(2)/Lai_Chung_Po_2015_JCE_2(2)_p8-17.pdf

Urbanska, K., & Guimond, S. (2018). Swaying to the extreme: Group relative deprivation predicts voting for an extreme right party in the French presidential election. *International Review of Social Psychology*, *31*(1). https://doi.org/10.5334/irsp.201

van Zomeren, M., Postmes, T., & Spears, R. (2008). Toward an integrative social identity model of collective action: A quantitative research synthesis of three socio-psychological perspectives. *Psychological Bulletin*, *134*(4), 504–535. https://doi.org/10.1037/0033-2909.134.4.504

Waldrop, M. M. (2021). News feature: Modeling the power of polarization. *Proceedings of the National Academy of Sciences of the United States of America*, *118*(37). https://doi.org/10.1073/pnas.2114484118

Westfall, J., Van Boven, L., Chambers, J. R., & Judd, C. M. (2015). Perceiving political polarization in the United States: Party identity strength and attitude extremity exacerbate the perceived partisan divide. *Perspectives on Psychological Science*, *10*(2), 145–158. https://doi.org/10.1177/1745691615569849

Wilson, J. P., & Hugenberg, K. (2010). When under threat, we all look the same: Distinctiveness threat induces ingroup homogeneity in face memory. *Journal of Experimental Social Psychology*, *46*(6), 1004–1010. https://doi.org/10.1016/j.jesp.2010.07.005

Wright, S., Taylor, D. M., & Moghaddam, F. M. (1990). Responding to membership in a disadvantaged group: From acceptance to collective protest. *Journal of Personality and Social Psychology*, *58*(6), 994–1003. https://www.researchgate.net/publication/38436432

Yustisia, W., Shadiqi, M. A., Milla, M. N., & Muluk, H. (2020). An investigation of an expanded encapsulate model of social identity in collective action (EMSICA) including perception of threat and intergroup contact to understand support for Islamist terrorism in Indonesia. *Asian Journal of Social Psychology*, *23*(1), 29–41. https://doi.org/10.1111/ajsp.12372

![cogitatio logo]

About the Authors

Rebekka Kesberg (PhD) is a senior research fellow in the School of Psychology at the University of Sussex, UK. Her research focuses on exploring justifications for political violence and its relationship with democratic values.

Allard R. Feddes (PhD) is an assistant professor at the Social Psychology Department at the University of Amsterdam, the Netherlands. He is interested in how group membership influences how we think, feel, and behave. Currently, his research focuses on the psychology of radicalization and terrorism, collective action, and discrimination.

Eva Vogel is a research assistant and master's student at the Social Psychology Department at the University of Amsterdam, the Netherlands. Currently, Eva uses her interdisciplinary background in politics, psychology, law, and economics, to better understand what drives (dis-)trust in political institutions and scientific authorities.

Bastiaan T. Rutjens (PhD) is an associate professor at the Psychology Research Institute of the University of Amsterdam, where he runs the Psychology of Science Lab.

cogitatio

Open Access Journal 🔓

(Micro)Identities in Flux: The Interplay of Polarization and Fragmentation in Polish and European Politics

Piotr Kłodkowski [1] 🆔, **Małgorzata Kossowska** [2] 🆔, and **Anna Siewierska** [3] 🆔

[1] Centre for Comparative Studies of Civilisations, Jagiellonian University, Poland
[2] Institute of Psychology, Jagiellonian University, Poland
[3] Institute of Political Science, University of Rzeszów, Poland

Correspondence: Piotr Kłodkowski (piotr.klodkowski@uj.edu.pl)

Abstract

This conceptual article interrogates the role of identity politics—amplified by political elites and magnified through social media—in destabilizing the foundations of liberal democracy by deepening ideological, social, and cultural divisions. It critically examines the dynamics of micro-identity formation, exploring their paradoxical capacity to either intensify polarization, erode democratic norms, and foster social fragmentation, or, conversely, to reinforce democratic norms. Through a case study of micro-narratives, such as the discourse surrounding the term "TERF" (trans-exclusionary radical feminist), the article highlights how the politicization of identity can fuel polarization and entrench societal fault lines. Equally, drawing on analyses of parliamentary elections in Poland, it illustrates how strategic integration of micro-identities into coalition-building efforts can yield positive outcomes, ultimately bolstering the democratic process. The article contends that safeguarding democratic norms necessitates the crafting of inclusive narratives that acknowledge micro-identities, uphold cultural and ideological pluralism, and reaffirm the core democratic principle of minority rights protection. In this context, particular attention is given to Polish PM Donald Tusk's post-2023 election discourse as an exemplar of an integrative narrative strategy.

Keywords

democratic norms; (de)polarization; European politics; micro-identities; Poland

---

1. Introduction

In liberal democracies, ideological diversity is not merely inevitable but foundational. Disagreements over economic policy, human rights, and the role of the state are intrinsic to the political process, with consensus-building serving as a pivotal mechanism in decision-making (e.g., Lipset, 1959). Unlike alternative democratic models—such as India's "majoritarian democracy" or the illiberal model espoused by Viktor Orbán in Hungary—liberal democracy is anchored in a core set of normative principles, including respect for democratic institutions, the peaceful transfer of power, civic engagement, freedom of expression and pluralism, the rule of law, and most crucially, mutual tolerance (Meyer, 2020). These democratic norms extend beyond the protection of individual civil liberties; they ensure legal protections for the free functioning of minorities, including religious, ethnic, and LGBTQ+ communities, as organized groups (Kołakowski, 2014; Safran, 2019). To maintain these norms, liberal democracies depend on independent judiciaries and strong civil institutions, which function as critical bulwarks against majoritarian encroachment, safeguarding the rights and participation of marginalized groups (Hilbink, 2012).

However, polarization—whether political, institutional, or affective—poses a profound challenge to these democratic principles, especially the protection of minority rights (Carothers & O'Donohue, 2019; Fukuyama, 2018). Much of this polarization can be traced back to processes of identity formation and the strategic exploitation of these identities by political elites through the use of social media (Baumer & Gold, 2015). Specifically, by targeting certain voter groups with tailored messages that align with their group's expectations, values, and norms, political actors tap into the strong sense of belonging individuals derive from their affiliations, often reinforcing normative pressures to conform (e.g., Bicchieri, 2005; Cialdini & Trost, 1998. Research has demonstrated that while such tactics can be highly effective in mobilizing support, they also may deepen societal divisions, exacerbating different types of inequalities, enflaming identity-based conflicts, and entrenching partisan hostility (Druckman, 2024; McCoy et al., 2018; Somer et al., 2023).

Although much previous research has explored the role of identity in broad categories—such as ethnicity, race, religion, or political affiliation—in fueling polarization (e.g., Castle & Stepp, 2021; West & Iyengar, 2022), the role of micro-identities (e.g., subgroups within larger categories) remains significantly underexplored. Moreover, actionable strategies for mitigating identity-driven polarization are neither well-documented nor systematically tested. Thus, this article seeks to address these gaps by investigating the role of micro-identities and their dual potential to either aggravate polarization or, when appropriately integrated into political decision-making, contribute to depolarization. We critically examine how micro-identities are frequently framed in opposition to one another, highlighting how political and cultural elites exploit these identities to deepen societal divisions. Using the case of trans-exclusionary radical feminist (TERF) narratives, we illustrate the ways in which identity-based conflicts are then mobilized to entrench polarization. However, we contend that when micro-identities are properly recognized and incorporated into political strategies, they can function as powerful mechanisms for depolarization and the reinforcement of liberal democratic norms. To support this claim, we present case studies from Poland's parliamentary elections, demonstrating how the strategic incorporation of micro-identities in coalition-building efforts can yield democratic outcomes. Next, we further analyze the construction of political narratives, distinguishing between two dominant types that shape the European political landscape. Central to our discussion is the imperative of crafting "inclusive narratives"—discourses that recognize and

respect the plurality of micro-identities within society. While such inclusive narratives may initially contribute to social fragmentation, we argue that, over time, they play a crucial role in mitigating polarization and sustaining the foundational democratic norms essential for the stable functioning of liberal democracy.

2. Polarized Social Identities Along Political Cleavages

The concept of political or ideological polarization refers to the widening divide between two major political factions on key societal issues (Fiorina & Abrams, 2008). This polarization manifests at two distinct levels: "elite polarization," characterized by increasing ideological divergence in the legislative behavior of elected officials, and "mass polarization," which reflects deepening ideological rifts within the general electorate. Lipset and Rokkan (1967), in their "cleavage theory," posit that political transformations are often driven by shifts in societal identities, since profound structural divisions within society become embedded in the political sphere through individual identity affiliations (Bartolini & Mair, 1990; Zollinger, 2024). Recent scholarship has increasingly underscored the centrality of social identities—how individuals perceive their connections to specific social groups—in shaping political attitudes and behavior (Achen & Bartels, 2017; Lipset & Rokkan, 1967; Young, 2023). Notably, some researchers argue that individuals often base their political choices not on policies or ideologies, but on their social identities—their deeply ingrained sense of who they are (Achen & Bartels, 2017; Young, 2023). This identity-centered approach has gained increasing traction, particularly with the rise of "affective polarization," wherein animosity toward political opponents overshadows substantive policy disagreements (Dimant & Kimbrough, 2024; Druckman et al., 2024; Iyengar et al., 2019; Sides et al., 2016).

Francis Fukuyama argues that the foundational principles of liberal democracy—grounded in a certain universal and ostensibly shared system of values and norms—are increasingly undermined by the rise of "identity politics" (Fukuyama, 2018). This phenomenon, driven by dynamic socio-political forces, appears to contribute to a weakening sense of community and may play a role in increasing social and political fragmentation (Fukuyama, 2018). Contemporary debates and scholarly discourse on identity politics in the US and Europe reflect the extent to which political, social, and cultural polarization is maintained, largely by political elites (Druckman, 2024; McCoy & Somer, 2019). This, in turn, translates into what scholars have termed a "democratic recession," characterized by the corrosion of long-established social norms once considered stable and immutable (Diamond, 2015). The strategic deployment of "identity politics" may serve as a mechanism for various political groups to bolster their influence and political power; through carefully crafted, polarizing narratives, these groups selectively appeal to specific components of these identities among their electorate or political constituencies. In explaining the process of polarization, Fukuyama (2018) highlights that left-leaning political movements in the US and Europe have played a key role in advancing multi-level identity-based frameworks that, while intended to empower marginalized communities, have simultaneously alienated other demographic groups. Specifically, he argues that identity-based advocacy has often prioritized the grievances of historically oppressed groups while sidelining populations traditionally associated with conservative values—such as individuals of European descent affiliated with Christianity, rural communities, and proponents of traditional family values and structures. This perceived exclusion is likely to fuel resentment among these groups, who feel unfairly blamed for societal discrimination. Consequently, these groups may begin to increasingly view themselves as victims of "liberal elites" in the US or "leftist elites" in Europe (Holmes, 2017).

In response to these perceived cultural and political shifts, traditional and more conservative groups who perceive themselves as marginalized could embrace countervailing identity frameworks emphasizing "familiarity" and resistance to perceived external threats. These perceived threats include mass migration from the Middle East, Africa, or South Asia, as well as the proliferation of multifaceted identities championed by many progressive European parties (Jansma, 2024; Ylä-Anttila, 2017). Within this context, such liberal frameworks might be seen as undermining cultural-national coherence, ultimately fueling political and affective polarization (Stichweh, 2021). This dynamic will possibly contribute to a resurgence of nationalist sentiments and a renewed attachment to dominant cultural identities, which, as Ozkirimli (2005) observes, often serves to rationalize and justify exclusionary policies and beliefs. Some empirical evidence supports these trends. Eurobarometer surveys (European Commission, 2021) reveal that family and nationality remain central to European self-identification, with 81% and 73% of respondents, respectively, citing these as core components of their identity.

Identity politics usually extends beyond individual and group affiliations, permeating institutional structures, thus posing a serious threat to the fundamental norms of liberal democracy. Institutional polarization has the potential to affect perceptions of impartiality and independence within state institutions, which may, in turn, influence public trust in mechanisms critical for upholding legal and social guarantees (Jezierska & Sörbom, 2021). In deeply polarized environments—including civil society organizations—state institutions will possibly become instruments of partisan agendas rather than neutral arbiters, further weakening democratic structures and trust in governance (Sörbom & Jezierska, 2023). The consequences of this institutional polarization also tend to spill over from civic institutions to others; as public trust in these organizations deteriorates, so too does trust in broader state institutions, triggering a wide-ranging systemic crisis (Druckman, 2024). A striking example of this trend emerged in Poland when, in December 2024, Polish Supreme Court Judge Piotr Prusinowski warned of a potential legitimacy crisis following the 2025 elections, cautioning that the country could face a scenario in which two rival presidents emerge, with neither faction accepting the other's authority ("Niebawem w Polsce może być dwóch prezydentów," 2024). This echoes the aftermath of the 2020 US elections, in which Donald Trump's supporters refused to recognize his defeat, alleging widespread electoral fraud and claiming the election was "stolen" (Viala-Gaudefroy, 2020). Such distrust in institutional neutrality—whether in the judiciary, the media, or civil society—might signal a dangerous acceleration of polarization that threatens democratic governance. As confidence in institutions wanes, the perceived legitimacy of democratic processes may weaken, increasing the likelihood of governance crises and heightened social conflict.

3. Identity Politics and Emerging Micro-Identities

Fukuyama (2018) conceptualizes identity politics through the lens of "isothymia"—a term he derives from the Greek *thymos*, denoting the intrinsic human desire for recognition in the eyes of others. This concept encapsulates a blend of pride, self-worth, and dignity, as well as an acute sensitivity to one's social positioning as an autonomous individual. He deems "isothymia," the struggle for recognition, to be the engine of world history, and a driving force behind contemporary identity politics, leading to the construction of distinct cultural group identities that diverge from the mainstream majority (see the description of the process of subgroup or dual identity formation in Verkuyten, 2013). These identity formations are exemplified by communities such as Roman-Catholic Latinos in the US, LGBTQ+ groups, and various Muslim communities across Europe—each shaped by distinct origins and ideological orientations.

A defining characteristic of modern identity politics is the increasing fragmentation of broad identity categories into more nuanced and distinct micro-identity groups, each with its own unique experiences, reflecting the phenomenon of "cross-cutting cleavages"—a process wherein individuals belonging to the same overarching identity group develop varied, divergent, and at times conflicting, allegiances (Dunning & Harrison, 2010; Root, 1996; Verkuyten, 2013).

This trend is particularly evident in the growing recognition of identities formed at the intersection of multiple axes of discrimination; for instance, the lived experience of a Black lesbian woman cannot be wholly subsumed under the broader categories of either race, gender, or sexual orientation alone; rather, it is impinged upon by the intersectionality of all three dimensions (Bernstein, 2005). In immigration, multiculturalism, and social integration contexts, individuals may hold complex or hybrid identities that bridge different cultural, ethnic, national, or social groups (Settles & Buchanan, 2014; Verkuyten, 2013). For example, dual (or multiple) identities (Settles & Buchanan, 2014; Verkuyten, 2013) reflect the simultaneous identification with two or more distinct social or cultural groups, typically involving a minority group and a larger, dominant group (e.g., Root, 1996). Representatives of these micro-groups increasingly advocate for the recognition of their full subjectivity within mainstream society, calling for a reinterpretation of existing social and legal norms that, in their view, have either historically marginalized them in the past, or continue to disadvantage them now as individuals and communities (Bernstein, 2005; Gitlin, 1995). Given that the protection of minority rights is a universally accepted norm in liberal democracy, such demands are often considered justified within democratic frameworks (Biró & Newman, 2022). Building on Fukuyama's (2018) work, a micro-identity—a type of subgroup identity (Rabinovich & Morton, 2011)—can be defined as a deeply subjective sense of belonging to an exceptionally specific and narrowly defined group. Members of these groups are bound by a shared identification with particular aspects of social reality and a strong belief in their distinctiveness (see also Verkuyten, 2013). In the digital realm, micro-identities often emerge as fluid and transient entities, characterized by distinct epistemic realities, unwavering internal support for their ideology and activities, and a strong sense of in-group self-determination (Lüders et al., 2022; see also van Zomeren et al., 2008). These identities are fragmented, context-dependent, and emotionally charged, coalescing around composite ideas and belief systems—features that set them apart from broader identity constructs (Gartenstein-Ross et al., 2023).

Emerging micro-identities and their corresponding micro-narratives are continuously evolving, embodying a sense of fluidity in which identities are dynamic and subject to change over time (Verkuyten, 2013; Verkuyten & Martinovic, 2012). Meanwhile, more established micro-identities persist, adapting and reshaping themselves in response to shifting social and political landscapes. This perpetual process reflects individuals' fundamental search for belonging, meaning, and purpose within their worldviews (for a review see Kruglanski et al., 2025). Importantly, micro-identities do not inherently preclude the coexistence of other identities or affiliations, including those rooted in religious, ethnic, and national affiliations (Verkuyten, 2013). For instance, an individual adhering to fundamentalist Islam in Europe may construct their identity based on Wahhabi ideology, viewing themselves as a "true Muslim." While this micro-identity, in theory, does not clash with the broader concept of Muslim identity, the majority of Muslims do not subscribe to this radical interpretation of Islam and can openly reject it. Nonetheless, this commonality with conventional religiosity does not deter its adherents from pursuing their own ideological objectives; in fact, they may simply perceive other Muslims as "unaware of their true identity" and work towards "raising their awareness" to a more "authentic" understanding of their faith (e.g., Firro, 2018; Rashid, 2003). Evidently, two parallel

processes are at play: the formation of a Muslim (minority) community across national borders (previously existing mainly in theory as the *umma*), and the simultaneous emergence of distinct micro-identities within that community itself (Kossowska et al., 2023).

One of the key drivers of social identity fragmentation is the increasingly segmented media landscape of recent decades (Chaffee & Metzger, 2001; Mancini, 2013; Van Aelst et al., 2017). The sheer proliferation in the number of media outlets has dispersed audiences across an ever-expanding array of platforms, many of which prioritize political agendas over journalistic objectivity (Benkler et al., 2018; Holt, 2016; Nygaard, 2019). Digital media further fast-tracks this process by fostering fluid and context-dependent modes of engagement, enabling individuals to navigate between discussions that emphasize different facets of their social identities (Cork et al., 2020; Hopthrow et al., 2020). We suggest that these shifting contexts may facilitate increasingly dynamic and fluid identification processes, encouraging further fragmentation of identities. Beyond media dynamics, fragmented identities may also stem from the attenuating ties between citizens and political parties. Over the past several decades, party identification has declined, while voter mobility and electoral volatility have risen (Bartolini & Mair, 1990; Dalton, 2018; Zollinger, 2024). Indeed, studies suggest that individuals' political attitudes are now more sophisticated and multifaceted than the rigid, simplistic cleavages reflected in conventional party politics (Caprara et al., 2006; Piurko et al., 2011).

In summary, while polarization remains a prominent concern in contemporary political discourse, the increasing specificity of identity categories, the growing fragmentation of media environments, and the erosion of stable partisan loyalties collectively indicate a broader trend towards social and political fragmentation. This shift may reflect the complexity of modern identity dynamics, where people's allegiances and identifications are fluid, multifaceted, and often at odds with established political and social structures. Therefore, the very existence of micro-identities in pluralistic societies should be taken into account by the political class both when crafting appropriate narratives for their voters and while forming suitable electoral coalitions. This consideration is particularly critical in an era of heightened cultural, institutional, and political polarization. A more finessed interpretation of micro-identities—and their strategic integration into political discourse and policymaking—holds the potential to mitigate polarization and contribute to a more effective process of depolarization, a subject we explore in further detail in the following sections.

3.1. Polarizing Potential of Micro-Identities: A TERF Case Study

The concept of micro-identities plays a fundamental role in the context of socio-political functioning, raising the question of just how "micro" these identities can actually be. Feminist identity, for instance, is often perceived—particularly in right-wing discourse—as a monolithic and highly cohesive movement. However, since at least 2008, a fierce internal dispute has fractured feminist communities, leading to the exclusion of certain feminists from mainstream discourse on women's rights. This rift became particularly visible with the emergence of the term "TERF," coined by feminist writer Viv Smythe to describe activists who reject the notion that transgender women are truly women, and who thus advocate for the exclusion of transgender women from women's spaces and organizations. Since its introduction, the label "TERF" has been widely used to describe feminist individuals whose opposition to legislative initiatives supporting transgender rights has led to their marginalization within feminist movements (Smythe, 2018).

The prominence of the TERF identity was significantly amplified by the British author J. K. Rowling, who has repeatedly expressed the view that completely disregarding biological gender is an abusive act and constitutes another form of discourse that aims to exclude women. She has articulated a position that acknowledges transgender rights while simultaneously asserting the distinctiveness of female lived experiences:

> I respect every trans person's right to live any way that feels authentic and comfortable to them. I'd march with you if you were discriminated against on the basis of being trans. At the same time, my life has been shaped by being female. I do not believe it's hateful to say so. (Rowling, 2020)

Despite this, many feminists, including Rebecca Solnit, reject Rowling's concerns, arguing that trans women do not pose a threat to cis women and that exclusionary rhetoric risks further marginalizing them (Solnit, 2020).

Rowling, however, like other feminists skeptical of transgender women, opposes the liberalization of laws concerning gender identity change, arguing that legislative reforms could be exploited by sexual predators to gain access to spaces designated for women. In Poland, criminologist Magdalena Grzyb of Jagiellonian University has voiced similar concerns. She holds the view that the recognition of *felt gender* as the sole criterion for legal gender change undermines feminist principles and risks exacerbating discrimination against women. Additionally, Grzyb has contended that trans individuals are not the most excluded group in Polish society, as groups such as the Roma and homeless populations experience even greater exclusion. These comments sparked intense controversy, leading to the cancellation of one of her university lectures following backlash from trans activists. The furor surrounding her remarks erupted on social media and in the pages of *Kultura Liberalna*, an influential Polish intellectual magazine, further illustrating the deeply polarized nature of contemporary feminist discourse (Grzyb, 2021). Initially, the term TERF was used specifically to describe feminists who distanced themselves from trans-inclusive feminism. However, it quickly took on additional connotations and, in some cases, was re-purposed into a political weapon wielded by exclusionary factions. As a result, the term has become a permanent fixture of contemporary feminist discourse, often invoked in ideological battles over gender identity and rights (Hines, 2017).

Rowling's influence, combined with feminist debates over trans rights, has reverberated in Poland, where right-wing media have distorted the conversation to serve their own agenda. Conservative outlets have convinced many Poles that the left seeks to erase the word "woman" from public discourse (Tomaszewski, 2021). They have not only cited Rowling but also Urszula Kuczyńska, a feminist activist and linguist affiliated with the Lewica (The Left) party, who—like Rowling—expressed concerns about replacing "woman" with terms such as "person with a uterus" (Romanowska, 2020). The backlash within progressive circles over her statement overshadowed the Lewica party's broader political and economic agenda, shifting public discourse almost exclusively toward LGBTQ+ issues. This shift likely contributed to Lewica's unexpectedly low support in the 2023 elections, as many traditionally left-wing voters—including segments of the working class—migrated toward Koalicja Obywatelska (Civic Coalition [KO]), a party that offered a broader ideological spectrum (Pawłowska & Korzeniowska, 2023). This case demonstrates how identity debates, when amplified and politicized, can reshape electoral outcomes—not only by fuelling internal divisions within progressive movements but also by providing right-wing narratives with a powerful wedge issue to exploit.

The TERF case exemplifies what can be an intrinsic feature of micro-identities: their tendency to be forged in opposition to other ideological narratives. In politicized contexts, micro-identities may emerge through

resistance, rebellion, and dissent, by defining themselves in contrast to narratives considered too mainstream, generic, or universalist. The TERF faction's critical view of trans women functioned as an internal pivot, setting them against even other competing micro-identities. In this sense, politicized micro-identities often derive their cohesion less from an intrinsic sense of belonging to a specific minority group and more from a perceived separateness from the majority (Gerbaudo et al., 2023). Unlike conventional social identities, which are typically rooted in self-perception and personal identification, micro-identities are primarily defined through external relational dynamics—shaped by opposition to other social groups, and the values, norms, and attitudes they represent (Gartenstein-Ross et al., 2023; Törnberg, 2022). However, it is important to recognize that not all micro-identities are defined by antagonism toward dominant socio-political frameworks. Some micro-identities adopt a more neutral stance, forming around common interests, cultural affiliations, or philosophical standpoints without necessarily positioning themselves in opposition to the political mainstream (Wachter et al., 2015).

3.2. Micro-Identities in Political Elections: A Case Study of Poland

Political, affective, and institutional polarization occurs within increasingly pluralistic yet deeply unequal societies, where political divisions frequently overlap with ideological, social, and religious divisions (Wilson et al., 2020). In any state, while opposition groups may share a common critique of the ruling powers and a critical view of the policies being proposed or implemented, they often diverge on numerous other issues. In Poland, opposition to the ruling Prawo i Sprawiedliwość (PiS; Law and Justice) government has encompassed not only left-wing groups but also conservatives who opposed the government's foreign policy—particularly its confrontational stance toward the European Union (EU)—and the politicization of state institutions, especially the judiciary. Consequently, building an effective democratic opposition required political actors to account for not only the overarching phenomenon of polarization but also for the existence of micro-identities, which may have their own political dimension (Paterson & Witzleb, 2020). In other words, greater party pluralism—meaning a larger number of diverse political groups or unique political offerings—and rising factionalism within major parties are not anomalies but rather natural consequences of an electorate whose ideological composition is becoming increasingly complex and fragmented, with burgeoning voter identities and growing ideological pluralism. The proliferation of political offerings in the marketplace of ideas reflects the corresponding diversification of voter identities and communities, necessitating a recalibration of political strategies to accommodate this new reality (Caprara & Vecchione, 2017). Accordingly, political actors and various ideologues must craft more inclusive and comprehensive narratives for their audiences that acknowledge cultural and ideological pluralism, while simultaneously constructing experimental yet viable post-election coalitions to secure power (Fortunato, 2021). This phenomenon was evident after Spain's 2019 general elections, when Prime Minister Pedro Sánchez formed a coalition between the Spanish Socialist Workers' Party (PSOE) and Unidas Podemos—an unprecedented experiment in coalition governance since the Second Spanish Republic. Four years later, Sánchez replicated this approach by forming another coalition government, this time with Sumar, a newly established left-wing political platform launched by Yolanda Díaz, with Unidas Podemos also participating in the new governing alliance (Field & Teruel, 2023). A similar dynamic played out in Poland three months later. The Polish case is particularly instructive because the opposition parties against the then-government, collectively referred to as the United Opposition, successfully ousted PiS, a party widely regarded in Europe as a populist right-wing grouping that, despite its defeat, continued to command significant electoral support. However, opposition leaders faced considerable strategic dilemmas in the run-up to the parliamentary elections.

Prior to the parliamentary elections in the autumn, two clashing electoral strategies emerged within the Polish opposition parties. Advocates of a unified opposition list, who rallied around the social media hashtag #silnirazem ("strongtogether"), argued that only a fully consolidated front—one that neutralized internal ideological divisions—could overcome the ruling United Right coalition, led by PiS, and take power in Poland (Przeworski, 2024). Conversely, critics of this approach contended that such enforced unity inherently excluded certain voters from the political discourse, i.e., those with specific priorities and expectations, including feminists and LGBTQ+ activists, on the one hand, and conservative opposition supporters on the other. The latter group rejected PiS's right-wing populism but remained wary of the dominant liberal-progressive orientation of Donald Tusk's KO. This fragmentation within the opposition created potential for cognitive dissonance among many voters, both those with more left-wing and liberal views, and those attached to traditionally conservative values, even potentially discouraging electoral participation altogether. In Table 1 we provide an overview of how different ideological groups and their micro-identities shaped the 2023 Polish election results, hoping that it helps non-experts understand the political landscape and coalition-building dynamics.

Ultimately, partly as a result of political calculations, and partly due to the individual ambitions of various party leaders, rather than coalescing into a single electoral bloc, opposition forces opted for a hybrid strategy. The opposition entered the elections as three separate, but strategically aligned, coalitions: KO, Trzecia Droga (Third Way; a center-right alliance between the Polish People's Party [PSL] and Poland 2050, a centrist party), and Lewica (a progressive-leftist electoral alliance). This multi-coalition approach proved successful, leading to a decisive electoral victory.

The decision to abandon a single opposition list in favor of an intermediary electoral strategy was probably the result of drawing the right conclusions from a United Surveys election poll (Żółciak & Osiecki, 2022) which indicated that, while 62% of opposition party supporters backed the idea of a unified electoral list, a substantial

Table 1. An overview of how different ideological groups and their micro-identities shaped the 2023 Polish election results.

Party	Political Spectrum	Social Values	Key Support Groups (Microidentities)
Lewica	Far-left to center-left	Socially liberal	Feminists, LGBTQ+ activists, climate change activists, socialist youth, progressive intellectuals
KO	Center-left to center	Socially liberal to moderate	Urban professionals, pro-EU liberals, business elites, progressive youth
Trzecia Droga	Center to center-right	Moderately conservative	Moderate conservatives, small-town business owners, pragmatic centrists
PiS	Right-wing conservative	Socially conservative	Rural conservatives, older voters, traditional Catholics, nationalist youth, anti-LGBTQ+ groups
Konfederacja	Far-right	Ultra-conservative and libertarian	Anti-tax libertarians, ultra-nationalists, young men disillusioned with mainstream politics

Source: Based on Centrum Badania Opinii Społecznej (2023).

minority (31%) expressed concerns that such a move would dissuade them from voting at all. The strategic flexibility of the opposition coalition ensured that a diverse electorate—dissatisfied with PiS rule but unwilling to compromise on ideological principles—could still find political representation without feeling forced into an ideologically monolithic bloc. In December 2023, Donald Tusk became prime minister and formed a new government incorporating all the major democratic opposition groups. While it remains a matter of speculation whether a single opposition list would have secured a comparable electoral outcome, the exceptionally high voter turnout suggests that the majority of voters dissatisfied with the rule of the United Right had found their political and ideological niche; in other words, a significant portion of the electorate was able to align with political options that resonated with their ideological and identity-based preferences without making too many concessions to their own beliefs (Jaworska-Surma et al., 2023). The possibility of choosing from a broader range of programs was particularly critical for younger voters, who tend to prioritize more granular political and cultural concerns that align with their emerging micro-identities.

The Polish case suggests that the opposition's electoral success was, at least in part, an inadvertent response to this increasingly relevant phenomenon of emerging micro-identities, which are making a tangible impact on the electoral process and, subsequently, on the policies implemented. However, while micro-identities reflect the diversity inherent in pluralistic democracies, their instrumentalization in politics carries risks. Liberal democracy is based on compromise, and without at least a minimalist but foundational political narrative, it is difficult to envision how a lasting parliamentary majority can be built. Micro-identities, in and of themselves, it must be emphasized, do not pose a direct threat to the stability of the state per se; rather, it is their strategic exploitation that creates a potential for disintegration and political volatility, making moderate majorities hostages to a vocal and active minority. The challenge, therefore, lies in preventing public discourse from becoming dominated by narrow, highly vocal groups whose demands, while legitimate, may disproportionately influence broader governance structures.

A potential solution lies in cultivating political communication centered on shared goals—however minimal those objectives may be. By fashioning unifying narratives that transcend factional divisions while still recognizing the legitimacy of micro-identities, political actors can mitigate polarization and create the groundwork for long-term democratic stability. This approach, if effectively applied, could work as a viable strategy for diminishing social and political polarization in the long term.

Despite its electoral success, the "coalition of many coalitions" faces a fundamental challenge: ideological cohesion. As Poland's ruling bloc consists of diverse factions with differing priorities, deep-seated divisions on sensitive issues like abortion and the role of the Church in public life may make it increasingly difficult to maintain unity. A striking example of this tension emerged when Minister of Culture and National Heritage Bartłomiej Sienkiewicz decided to cut state subsidies for religious publications, publicly declaring that "the mission of the state is not to spread faith and salvation" ("This is a secular state," 2024). His stance aligns with the secular vision championed by Civic Platform (PO, which is the strongest party within KO) and Lewica, who advocate for a clearer separation between Church and State. However, a different perspective was offered by Deputy Prime Minister Władysław Kosiniak-Kamysz, leader of the coalition conservative PSL. While he supports limiting the Church's direct involvement in politics, he insists that policy should still be grounded in Christian values ("Kosiniak-Kamysz: PSL konsekwentnie broni wartości chrześcijańskich," 2019). This debate unfolds against the backdrop of Poland's rapid secularization, particularly among younger generations. A 2018 Pew Research Center report identified Poland as one of the fastest-secularizing countries globally, based on

the widening gap in religiosity between younger and older demographics ("Proportion of Catholics in Poland," 2023; "The age gap in religion," 2018). The 2021 National Census further underscores this shift, revealing that 71.3% of Poles identified as Roman Catholic, a sharp decline from 87.6% in 2011 ("Coraz mniej katolików," 2023; Główny Urząd Statystyczny, 2021). These trends suggest that proposals for further secularization—championed by PO and Lewica—are likely to resonate with younger voters. However, the coalition's ideological diversity means that secularization reforms will likely be tempered by compromise. With both PO and PSL operating within the ruling coalition, any major push toward a more secular state is expected to be moderate, carefully balancing progressive aspirations with conservative sensibilities.

The key political question is whether this cautious, incremental approach will prove sufficient to keep younger, secular-minded voters engaged. If reforms fall short of their expectations, disillusionment could drive voter apathy, weakening support for the coalition in future elections. Ultimately, much will depend on Prime Minister Donald Tusk's ability to maintain credibility, sustain coalition unity, and craft a compelling vision for governing a deeply divided society in an era of growing political polarization.

4. Searching for Common Ground Addressing Micro-Identifications

The role of "grand political narratives" is crucial in shaping socio-political dynamics. These overarching narratives function as an ideological backdrop against which more personal personalized stories can be crafted, and ideological identities developed and refined (Mayer, 2014). The process of identity formation is continuous but labile, susceptible to evolving social contexts and technological developments. The increasing individualization of society—intensified by the logic of algorithm-driven digital environments—stimulates the birth of additional micro-identities (Kossowska et al., 2023), potentially accelerating disintegration processes with far-reaching and dramatic repercussions.

In this fragmented landscape, the urgent task for a responsible political class is to construct a maximally inclusive narrative that resonates across a broad spectrum of the electorate. Such an endeavor is complex, as it must accommodate the diverse needs of an increasingly pluralistic society. Historical precedents provide valuable insights; in Poland, the "Solidarity" movement of the 1980s stands out as a unifying narrative that successfully brought together individuals from varied social and ideological backgrounds (Brier, 2021). The political context of that time was pivotal, with a clearly defined opponent—the oppressive and economically inept state—with which negotiations could be conducted, leading to a peaceful compromise. By way of contrast, the predicament democratic countries find themselves in today seems to be much thornier (Foret, 2024).

We identify two predominant narratives existing in the cultural-political space of Europe. The first, tentatively labeled "cultural-liberal" and "pan-European," embodies a form of universal social norms, with similar characteristics across most countries in Western, Northern, Central, and Southern Europe. As an overarching framework, it puts an emphasis on the secular nature of the state (for example, opposing the overt display of religious symbols in public spaces), and is liberal on moral issues—strongly supporting reproductive rights and same-sex partnerships, and open (or at least not vehemently opposed) to immigration including ethnic and religious groups from outside the Western sphere. It thus envisages a multicultural society while being skeptical of "national values," or even the concept of a "nation-state" grounded solely on ethnic principles (Duranti, 2017). Political parties or a coalition of parties, such as KO in

Poland, the Social Democratic Party (SPD) of Germany, and La République En Marche! in France, conform to this narrative. Overall, this narrative advocates for a high level of income redistribution, criticizes global capitalism and economic neoliberalism, and places various interpretations of equality and social justice on a pedestal. Within this framework, the EU is envisioned as a cohesive entity built on shared ideological foundations, including a commitment to moral diversity, tolerance, and the principle of secular governance. For all its celebration of a diversity of traditions, languages, and customs, this narrative comes with a caveat—not always explicitly stated—that this multiplicity falls within "European moral values and social norms." These norms include tolerance for various minorities within the framework of multiculturalism policy and/or the acceptance of secular state principles (Triandafyllidou et al., 2011). The intellectual underpinnings of this narrative can be traced back to theorists like Ernest Gellner, Eric Hobsbawm, or Benedict Anderson, who describe how modern nations can be engineered into a deliberate "construct" rather than, for example, the culmination of organic, historical processes of development of a national identity (Anderson, 2006; Gellner, 2009; Hobsbawm, 1992). From this perspective, it is thought to be perfectly feasible to "overcome" the risk of nationalism and "reconstruct" societies according to more liberal formulas. Even if not always explicitly articulated, the aspiration for a pan-European identity serves as a crucial reference point, despite the controversies and doubts it raises (Applebaum, 2020). The responsibility for advancing this vision rests with the political and intellectual elites across the Old Continent, who are fully cognizant that only a more unified and integrated Europe can effectively respond to external threats (Hampshire, 2014).

On the other hand, the second narrative—which we term "national conservative"—accepts, though not always without contention, the functioning of the EU in its current political form while strongly contesting the current immigration policy. Central to this narrative is its concern about unchecked immigration, with the result that this issue has become dominant in domestic political struggles. Limiting migration, particularly from Muslim-majority countries, and, in the case of Western European countries, from the so-called "new EU member states in the East" is a major strand in this narrative, stemming from the fear that excessive cultural diversity (especially involving non-European elements) could cause the cultural and social coherence of the state to crumble, potentially leading to soaring crime or even anarchy in certain areas. Not surprisingly, this perspective emphasizes "national security" and "law and order," particularly in response to various forms of social unrest (including riots) in neighborhoods predominantly inhabited by immigrant communities and their descendants (Murray, 2018). The issue of migration and security is intricately linked to individual states' attitudes toward the concept of the EU itself, symbolized by "Brussels." The "national conservative" narrative reflects growing skepticism, not only toward any form of political centralization at the pan-European level but also toward even moderate proposals to streamline collective actions across the Union. The reason for this is that it is believed these could diminish the role of individual member states, for example, by weakening the veto of a single country, which might thereby paralyze a community's ability to act in some areas of national interest (Caiani et al., 2024). Hence, "Brussels" is viewed, if not as an enemy, then at least as an adversary that must be dealt with firmly. Thus, there is a strong emphasis on the fundamental principle of subsidiarity, which lies at the core of the EU, even if is not always positively received within the "socio-liberal" narrative. This "national conservative" narrative draws, to some extent, on the French Annales School, particularly its prominent representative, Fernand Braudel, and his concept of *longue durée*, which emphasizes the enduring nature of cultural and national identities (Braudel, 1982, 1995). While elements of national constructivism are incorporated into the narrative itself, it primarily intends to promote a vision of national identity as a product of natural, long-term development, resistant to external influences (Duranti, 2017).

Both narratives are, at their core, political constructs, rarely manifesting in pure ideological form within the social realities of individual European countries. Nonetheless, they serve as broad narrative frameworks from which more specific ideological messages are derived (Mayer, 2014). Depending on the country, issues such as migration, the role of religion in public life, or economic policies might be emphasized differently. However, this ideological dichotomy—more or less refined across various parts of the Union—seems to delineate the political contours of Europe. Even if we view these two narratives as representing extremes on a broad ideological continuum, with individual parties positioning themselves closer to or further from these extreme points, a fundamental ideological division is clearly apparent. One might postulate that the field of ideological compromise between these narratives—and thus between individual parties or party coalitions representing them—appears to be narrowing, with the extremes becoming more pronounced. This phenomenon is perhaps most evident in the US, where the ideological overlap between the two major parties has nearly vanished, effectively eliminating the once-present categories of "liberal Republicans" and "conservative Democrats" (Kleinfeld, 2021).

This growing polarization of narratives implies that individuals increasingly feel compelled to either identify with one of these narratives, even if they do not fully endorse all aspects, or to search out alternative ideological positions outside the established continuum. This trend contributes to the proliferation of narratives within the existing ideological continuum or to the creation of various narratives that challenge the existing mainstream political system reflected in the continuum. Notably, research conducted in Europe and the US indicates that a segment of the population remains unaligned with either dominant narrative, seeking instead a new ideological offering which may not be within the continuum. (Van Hauwaert & Van Kessel, 2018). The evident recent electoral success in Poland of a coalition comprising multiple party alliances, each showcasing distinct political narratives, illustrates this hypothesis. However, the sustainability of such a coalition hinges on constructing a unifying narrative—one with a minimal effective dose of ideological content while addressing key issues, thereby realistically constituting the lowest common denominator, palatable to the majority of voters (Capano et al., 2023). In Poland, for example, this involves the issues of national security (amidst the ongoing war in Ukraine and Russia's belligerent stance) and the preservation of a just rule of law. Ideological differences within such a coalition are natural, and heated negotiations and a search for common ground comprise the attendant norms of such a democratic system.

At the same time, the formation of a parliamentary majority of this kind entails risks, including the potential overreach of power under the guise of democratic legitimacy. A democratic litmus test for such a parliamentary majority lies in its adherence to civil liberties and minority rights, even when these principles are subject to divergent interpretations. As Leszek Kołakowski noted years ago:

> Tocqueville...saw both the advantages and dangers of various political systems, and his famous remarks on democratic tyranny, the tyranny of the majority, have certainly not aged....For this reason, civil liberties often seem more important to us than the rule of the majority, considering that majority rule can degenerate into terrible despotism, and public life can sometimes be tolerable without majority rule. (Kołakowski, 2014, p.190)

Compromise, though often at the expense of the idealized ideological purity of each side, must prioritize respect for civil liberties and minority rights as the cornerstone of any political agreement within a coalition.

Constructed effectively, such a narrative can provide a counterbalance to opposing narratives that may lean toward more authoritarian governance models.

5. An Attempt to Construct an Inclusive Narrative in Poland After the 2023 Elections

As the leader of a multiparty and ideologically pluralistic coalition, Polish Prime Minister Donald Tusk attempted to construct an inclusive narrative that draws from both cultural-liberal/pan-European elements as well as conservative-traditionalist narratives. His approach has been to strive to apply the lowest common denominator—a shared set of priorities capable of resonating with the widest possible electorate, regardless of their ideological leanings. In many ways, this strategy mirrors the "coalition of many coalitions" model that emerged in Poland prior to the 2023 elections, blending liberal-leftist themes with more conservative appeals to national traditions (Buras, 2024). As prime minister of a coalition government, Tusk strongly emphasizes the issue of security, positioning it as a unifying concern not only for Poland but for Europe as a whole. This focus is particularly germane in the context of Poland's presidency of the Council of the EU (January–June 2025), a period during which security and geopolitical stability are expected to be dominant themes. The ongoing Russian invasion of Ukraine, coupled with the hybrid war tactics initiated by Russia and Belarus, is presented not only as a direct threat to Central and Eastern Europe but to the EU as a whole ("Donald Tusk: Musimy bronić naszych granic," 2024). Additionally, the outcome of the 2024 US presidential elections is anticipated to have significant ramifications for the EU's geopolitical positioning, even affecting the functioning of the EU itself. In his key address to the European Parliament, Tusk underscored both security issues and the values of liberal democracy, including minority rights. While drawing from the resources of both narratives analyzed here, he deliberately avoided resorting to the extreme nationalist interpretations often associated with security discourse:

> Security is extremely important in the day-to-day lives of all Europeans. Let me tell you this, if we believe that a change of presidency, of administration in the US is a leap into the unknown, remember this: Europe has never been afraid to leap into the unknown. I listen to other world leaders who remember their moments of greatness....I would like to tell you that you can tackle the problem of illegal migration and problems with internal security in Europe without nationalistic and xenophobic slogans. I would like us in this domain, just as with external security, to favor traditional democracy based upon traditional democratic values such as individual freedoms, respect for minority rights, freedom of speech, rule of law, public life free from corruption, politics free from the omnipotence of oligarchs. I would like all those who profess these values to show their power and force when we need to defend our borders and our security. (Tusk, 2025)

Delivered in Polish and translated simultaneously, this speech was directed not only at the members of the European Parliament but also at Tusk's domestic electorate ("Donald Tusk w Parlamencie Europejskim," 2025). By also briefly alluding to "Solidarity"—referencing an inclusive narrative that united diverse factions in the late 1980s—he sought to reinforce a broader message of political unity (Tusk, 2025). In national addresses, Tusk has further linked security concerns to the defense of freedom and, by extension, to the foundational principles of the EU. Speaking in Warsaw at the official inauguration of Poland's Council of the EU presidency, he stated: "If Europe is powerless, it will not survive. Let's do everything so that Europe and Poland do not have to pay the highest price for freedom, for strength, for sovereignty. Let's do everything to make Europe strong again" (Strzelecki et al., 2025).

These key speeches represent an attempt to construct a political message that appeals to a broad political and social spectrum of society, particularly in a pluralistic landscape where micro-identities play an increasingly influential role. However, the effectiveness of such a narrative is also inextricably linked to the credibility of the narrator. An empathetic mainstream politician can craft a message that, in theory, should be widely acceptable, but its reception will ultimately be shaped by public perceptions of the speaker's trustworthiness and political effectiveness in achieving their objectives (Capano et al., 2023). Consequently, criticisms of various aspects of Tusk's constructed narrative—whether from political opponents or disillusioned voters—may stem less from the content of his narrative and more from preexisting attitudes towards him as a political figure, despite his best intentions to construct the most inclusive narrative possible. Ultimately, the true test of both the narrator and the narrative will be the next parliamentary elections in Poland, which will determine whether this strategic attempt at inclusivity has successfully resonated with the electorate.

6. Conclusions

The role of micro-identities in processes of political polarization and depolarization remains insufficiently understood and warrants further study. However, its implications extend far beyond theoretical debates, carrying profound consequences for social cohesion, political stability, and governance within the Euro-Atlantic sphere and beyond. This topic is of interest not only to academics but also to individuals involved in various levels of politics. The issue is not merely one of social coherence and minimal cultural harmony within any particular state but of practical political strategy—particularly concerning how electoral decisions intersect with the self-definition of micro-identities across various segments of the electorate. As voter identities become increasingly pluralistic and specialized in terms of their increasingly narrow interests, these may not always align with mainstream narratives. Hence, a critical task for political elites in the coming years will be to construct more inclusive narratives that account for the dynamism of micro-identities in increasingly heterogeneous societies and uphold the norms of liberal democracy (primarily ensuring the protection of minority rights). If successfully delivered, such narratives could serve as an effective mechanism for mitigating polarization, and fostering a more stable and cohesive political order. This will likely be among the most pressing and complex responsibilities facing democratic leadership in the 21st century.

It is difficult to clearly assess how the examples of building ideologically diverse coalitions in Poland and Spain can be applied to other European countries. Nevertheless, it is worth proposing a working hypothesis that, in many EU countries, the principles of "ideological purity" and the promotion of carefully designed political platforms by individual parties will be replaced by the principle of creating effective coalitions composed of many parties, not necessarily closely aligned in terms of governance philosophy and ideological message. On one hand, this would reflect the growing ideological pluralism of voters and allow for the seizure of power in the state, but on the other, it would pose a significant challenge in terms of the political coherence of such a coalition, and thus its effectiveness in governance.

It is important to highlight the example of Germany, where two ideologically different parties, the CDU/CSU (Christlich Demokratische Union/Christlich-Sociale Union, an alliance of two parties) and the SPD, have managed, despite challenges, to form a credible coalition (Große Koalition) that has governed for many years. They are prepared to form the same coalition again (Martin, 2025). It is therefore possible that the new paradigm of building a moderate and inclusive "coalition of many coalitions" with its own narrative, aiming to

weaken radical and populist parties, will gain increasing acceptance in Europe. The coming years will reveal whether Europeans are ready for such solutions.

Acknowledgments

Funding

Conflict of Interests

The authors declare no conflict of interests.

References

Achen, C. H., & Bartels, L. M. (2017). *Democracy for realists: Why elections do not produce responsive government*. Princeton University Press.

Anderson, B. (2006). *Imagined communities: Reflections on the origin and spread of nationalism*. Verso.

Applebaum, A. (2020). *Twilight of democracy: The seductive lure of authoritarianism*. Doubleday.

Bartolini, S., & Mair, P. (1990). Policy competition, spatial distance and electoral instability. *West European Politics*, *13*(4), 1–16.

Baumer, D. C., & Gold, H. J. (2015). *Parties, polarization and democracy in the United States*. Taylor & Francis.

Benkler, Y. (2018). *Network propaganda: Manipulation, disinformation, and radicalization in American politics*. Oxford University Press.

Bernstein, M. (2005). Identity politics. *Annual Review of Sociology*, *31*, 47–74.

Bicchieri, C. (2005). *The grammar of society: The nature and dynamics of social norms*. Cambridge University Press.

Biró, A.-M., & Newman, D. (2022). *Minority rights and liberal democratic insecurities. The challenge of unstable orders*. Taylor & Francis.

Braudel, F. (1982). *On history*. University of Chicago Press.

Braudel, F. (1995). *A history of civilizations*. Penguin Books.

Brier, R. (2021). *Poland's solidarity movement and the global politics of human rights*. Cambridge University Press.

Buras, P. (2024, October 25). Tusk's asylum coup: Breaking away from populist narratives. *European Council on Foreign Relations*. https://ecfr.eu/article/tusks-asylum-coup-is-a-break-from-populism-possible

Caiani, M., Carlotti, B., Lovec, M., Wincławska, M., Kočan, F., & Balcer, A. (2024). *European narratives and Euroscepticism in the Western Balkans and the EU*. Taylor & Francis.

Capano, G., Galanti, M. T., & Barbato, G. (2023). When the political leader is the narrator: The political and policy dimensions of narratives. *Policy Sciences*, *56*, 233–265.

Caprara, G. V., Schwartz, S., Capanna, C., Vecchione, M., & Barbaranelli, C. (2006). Personality and politics: Values, traits, and political choice. *Political Psychology*, *27*(1), 1–28. https://doi.org/10.1111/j.1467-9221.2006.00447

Caprara, G. V., & Vecchione, M. (2017). *Personalizing politics and realizing democracy*. Oxford University Press.

Carothers, T., & O'Donohue, A. (Eds.). (2019). *Democracies divided. The global challenge of political polarization.* Brookings Institution Press.

Castle, J. J., & Stepp, K. K. (2021). Partisanship, religion, and issue polarization in the United States: A reassessment. *Political Behavior*, *43*, 1311–1335.

Centrum Badania Opinii Społecznej**.** (2023). *Zachowania prospołeczne i zaangażowanie obywatelskie Polaków* (Research Announcment No. 98/2023). https://www.cbos.pl/SPISKOM.POL/2023/K_098_23.PDF

Chaffee, S. H., & Metzger, M. J. (2001). The end of mass communication? *Mass Communication & Society*, *4*, 365–379.

Cialdini, R. B., & Trost, M. R. (1998). Social influence: Social norms, conformity and compliance. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *The handbook of social psychology* (4th ed., Vol. 2, pp. 151–192). McGraw-Hill.

Coraz mniej katolików, Polska szybko się sekularyzuje—najnowsze dane Spisu Powszechnego. (2023, September 29). *Oko Press*. https://oko.press/spis-powszechny-katolicy

Cork, A., Everson, R., Levine, M., & Koschate, M. (2020). Using computational techniques to study social influence online. *Group Processes & Intergroup Relations*, *23*, 808–826. https://doi.org/ 10.1177/1368430 220937354

Dalton, R. J. (2018). *Citizen politics: Public opinion and political parties in advanced industrial democracies.* CQ Press.

Diamond, L. (2015). Facing up to the democratic recession. *Journal of Democracy*, *26*(1), 141–155.

Dimant, E., & Kimbrough, E. O. (2024). Polarization in multidisciplinary perspective. *PNAS Nexus*, *3*(10), Article pgae425.

Donald Tusk w Parlamencie Europejskim. "Jeszcze Europa nie zginęła." (2025, January 22). *Onet*. https://wiadomosci.onet.pl/swiat/donald-tusk-w-parlamencie-europejskim-jeszcze-europa-nie-zginela/r48xb8k

Donald Tusk: Musimy bronić naszych granic. (2024, June 10). *Polska Zbrojna*. https://polska-zbrojna.pl/home/videoshow/1572?t=Donald-Tusk-Musimy-bronic-naszych-granic

Druckman, J. N. (2024). How to study democratic backsliding. *Political Psychology*, *45*, 3–42.

Druckman, J. N., Klar, S., Krupnikov, Y., Levendusky, M., & Ryan, J. B. (2024). *Partisan hostility and American democracy: Explaining political divides.* University of Chicago Press.

Dunning, T., & Harrison, L. (2010). Cross-cutting cleavages and ethnic voting: An experimental study of cousinage in Mali. *American Political Science Review*, *104*, 21–39.

Duranti, M. (2017). *The conservative human rights revolution. European identity, transnational politics, and the origins of the European Convention.* Oxford University Press.

European Commission. (2021). *Standard Eurobarometer 95.* https://europa.eu/eurobarometer/surveys/detail/2532

Field, B. N., & Teruel, J. R. (2023, December 5). Another leftist coalition government in Spain, though just barely. *Australian Outlook*. https://www.internationalaffairs.org.au/australianoutlook/another-leftist-coalition-government-in-spain-though-just-barely

Fiorina, M. P., & Abrams, S. J. (2008). Political polarization in the American public. *Annual Review of Political Science*, *11*, 563–588.

Firro, T. (2018). *Wahhabism and the rise of the House of Saud.* Sussex Academic Press.

Foret, F. (2024). *The European Union in search of narratives. Disenchanted Europe?* Taylor & Francis.

Fortunato, D. (2021). *The cycle of coalition. How parties and voters interact under coalition governance.* Cambridge University Press

Fukuyama, F. (2018). *Identity. The demand for dignity and the politics of resentment.* Farrar, Straus and Giroux.

Gartenstein-Ross, D., Zammit, A., Chace-Donahue, E., & Urban, M. (2023). Composite violent extremism: Conceptualizing attackers who increasingly challenge traditional categories of terrorism. *Studies in Conflict & Terrorism*. Advance online publication. https://doi.org/10.1080/1057610X.2023.2194133

Gellner, E. (2009). *Nations and nationalism*. Cornell University Press.

Gerbaudo, P., De Falco, C. C., Giorgi, G., Keeling, S., Murolo, A., & Nunziata, F. (2023). Angry posts mobilize: Emotional communication and online mobilization in the Facebook pages of Western European right-wing populist leaders. *Social Media + Society*, *9*(1). https://doi.org/10.1177/20563051231163327

Gitlin, T. (1995). *Twilight of common dreams: Why America is wracked by culture wars*. Metropolitan Books.

Główny Urząd Statystyczny. (2021). *NSP 2021*. https://stat.gov.pl/spisy-powszechne/nsp-2021

Grzyb, M. (2021, May 24). Transpłciowość, kapitalizm i ostracyzm. Odpowiedź Stanisławowi Krawczykowi. *Kultura Liberalna*. https://kulturaliberalna.pl/2021/05/24/magdalena-grzyb-odpowiedz-polemika-stanislaw-krawczyk-transplciowosc

Hampshire, J. (2014). *The politics of immigration. Contradictions of the liberal state*. Polity Press.

Hilbink, L. (2012). The origins of positive judicial independence. *World Politics*, *64*, 587–621.

Hines, S. (2017). The feminist frontier: On trans and feminism. *Journal of Gender Studies*, *28*(2), 145–157. https://doi.org/10.1080/09589236.2017.1411791

Hobsbawm, E. J. (1992). *Nations and nationalism since 1780: Programme, myth, reality*. Cambridge University Press.

Holmes, K. (2017). *The closing of the liberal mind: How groupthink and intolerance define the left*. Encounter Books.

Holt, D. (2016). Branding in the age of social media. *Harvard Business Review*, *94*, 40–50.

Hopthrow, T., Smith, L. G., & Levine, M. (2020). Rethinking the group: Group processes in the digital age. *Group Processes & Intergroup Relations*, *23*, 801–807.

Iyengar, S., Lelkes, Y., Levendusky, M., Malhotra, N., & Westwood, S. J. (2019). The origins and consequences of affective polarization in the United States. *Annual Review of Political Science*, *22*, 129–146.

Jansma, J. (2024). Our culture is best! Populist engagement with culture in the construction of the populist myth. *Journal of Intercultural Studies*. Advance online publication. https://doi.org./10.1080/07256868.2024.2359976

Jaworska-Surma, A., Bendyk, E., & Gutkowski, S. (2023). *Fenomen wyborczej mobilizacji. Przyczyny rekordowej frekwencji podczas wyborów parlamentarnych 2023—wnioski z badań*. Fundacja Batorego. https://www.batory.org.pl/wp-content/uploads/2024/01/Fenomen.wyborczej.moblizacji_RAPORT.pdf

Jezierska, K., & Sörbom, A. (2021). Proximity and distance: Think tanks handling the independence paradox. *Governance*, *34*, 395–411.

Kleinfeld, R. (2021). The rise of political violence in the United States. *Journal of Democracy*, *32*(4), 160–176.

Kołakowski, L. (2014). *Niepewność epoki demokracji*. Znak.

Kosiniak-Kamysz: PSL konsekwentnie broni wartości chrześcijańskich. Od 124 lat. (2019, May 4). *Rzeczpospolita*. https://www.rp.pl/polityka/art9359021-kosiniak-kamysz-psl-konsekwentnie-broni-wartosci-chrzescijanskich-od-124-lat

Kossowska, M., Kłodkowski, P., Siewierska-Chmaj, A., Guinote, A., Kessels, U., Moyano, M., & Strömbäck, J. (2023). Internet-based micro-identities as a driver of societal disintegration. *Humanities & Social Sciences Communications*, *10*, Article 955. https://doi.org/10.1057/s41599-023-02441-z

Kruglanski, A. W., Prilleltensky, I., & Amiram, R. (2025). *The Routledge international handbook of human significance and mattering*. Routledge.

Lipset, S. M. (1959). Some social requisites of democracy: Economic development and political legitimacy. *American Political Science Review*, *53*, 69–105.

Lipset, S. M., & Rokkan, S. (1967). *Cleavage structures, party systems, and voter alignments: An introduction* (Vol. 2). Free Press.

Lüders, A., Dinkelberg, A., & Quayle, M. (2022). Becoming 'us' in digital spaces: How online users creatively and strategically exploit social media affordances to build up social identity. *Acta Psychologica, 228*, Article 103643.

Mancini, P. (2013). The Italian public sphere: A case of dramatized polarization. *Journal of Modern Italian Studies, 18*(3). https://doi.org/10.1080/1354571X.2013.780348

Martin, N. (2025, March 8). Germany: CDU/CSU and SPD closer to coalition. *DW.* https://www.dw.com/en/germany-cdu-csu-and-spd-closer-to-coalition/a-71868076

Mayer, F. W. (2014). *Narrative politics. Stories and collective action.* Oxford University Press.

McCoy, J., Rahman, T., & Somer, M. (2018). Polarization and the global crisis of democracy: Common patterns, dynamics, and pernicious consequences for democratic polities. *American Behavioral Scientist, 62*(1), 16–42.

McCoy, J., & Somer, M. (2019). Toward a theory of pernicious polarization and how it harms democracies: Comparative evidence and possible remedies. *The Annals of the American Academy of Political and Social Science, 681*(1), 234–271.

Meyer, M. (2020). *Liberal democracy. Prosperity through freedom.* Springer.

Murray, D. (2018). *The strange death of Europe. Immigration, identity, Islam.* Bloomsbury Publishing.

Niebawem w Polsce może być dwóch prezydentów. Tak uważa sędzia SN. (2024, December 27). *Business Insider.* https://businessinsider.com.pl/wiadomosci/niebawem-w-polsce-moze-byc-dwoch-prezydentow-tak-uwaza-sedzia-sn/78r457f

Nygaard, S. (2019). The appearance of objectivity: How immigration-critical alternative media report the news. *Journalism Practice, 13*(10), 1147–1163.

Ozkirimli, U. (2005). *Contemporary debates on nationalism: A critical engagement.* Palgrave Macmillan.

Paterson, M., & Witzleb, N. (2020). *Attorney-General's Department review of the Privacy Act 1988.* Monash University.

Pawłowska, D., & Korzeniowska, K. (2023, October 23). Wybory 2023. Mobilizacja, przepływy elektoratów i nowi wyborcy. *Gazeta Wyborcza.* https://biqdata.wyborcza.pl/biqdata/7,159116,30309159,wybory-2023-mobilizacja-przeplywy-elektoratow-i-nowi-wyborcy.html

Piurko, Y., Schwartz, S. H., & Davidov, E. (2011). Basic personal values and the meaning of left–right political orientations in 20 countries. *Political Psychology, 32*(4), 537–561.

Proportion of Catholics in Poland falls to 71%, new census data show. (2023, September 29). *Notes from Poland.* https://notesfrompoland.com/2023/09/29/proportion-of-catholics-in-poland-falls-to-71-new-census-data-show

Przeworski, A. (2024). Who decides what is democratic? *Journal of Democracy, 35*(3), 5–16.

Rabinovich, A., & Morton, T. A. (2011). Subgroup identities as a key to cooperation within large social groups. *British Journal of Social Psychology, 50*(1), 36–51.

Rashid, A. (2003). *The rise of militant Islam in Central Asia.* Yale University Press.

Romanowska, K. (2020, December 12). Mam wątpliwości, czy zamiana "kobiet" na "osoby z macicami" to postawa lewicowa. *Wysokie Obcasy.* https://www.wysokieobcasy.pl/wysokie-obcasy/7,163229,26597103,kobiety-w-polsce-jeszcze-nie-zajely-miejsca-na-scenie.html

Root, M. P. P. (1996). *The multiracial experience: Racial borders as the new frontier.* Sage.

Rowling, J. K. [@jk_rowling]. (2020, June 7). *I respect every trans person's right to live any way that feels authentic and comfortable to them* [Post]. X. https://x.com/jk_rowling/status/1269407862234775552

Safran, W. (2019). Civil liberties in democracies: Constitutional norms, practices, and problems of comparison. In V. P. Nanda (Ed.), *Global human rights* (pp. 195–210). Routledge.

Settles, I. H., & Buchanan, N. T. (2014). Multiple groups, multiple identities, and intersectionality. In V. Benet-Martínez & Y. Hong (Eds.), *The Oxford handbook of multicultural identity* (pp. 160–180). Oxford University Press.

Sides, J., Tesler, M., & Vavreck, L. (2016). The electoral landscape of 2016. *The Annals of the American Academy of Political and Social Science, 667*(1), 50–71.

Smythe, V. (2018, November 28). I'm credited with having coined the word 'Terf.' Here's how it happened. *The Guardian*. https://www.pressreader.com/australia/the-guardian-australia/20181130/281814284925774

Solnit, R. (2020, August 10). Trans women pose no threat to cis women, but we pose a threat to them if we make them outcasts. *The Guardian*. https://www.theguardian.com/commentisfree/2020/aug/10/trans-rights-feminist-letter-rebecca-solnit

Somer, M., McCoy, J. L., & Luke, R. E. (2023). Pernicious polarization, autocratization and opposition strategies. *Democratization, 28*(5), 929–948. https://doi.org/10.1080/13510347.2020.1865316

Sörbom, A., & Jezierska, K. (2023). Social capital and polarization: The case of Polish think tanks. *Journal of Civil Society, 19*(4), 347–365.

Stichweh, R. (2021). How do divided societies come about? Persistent inequalities, pervasive asymmetrical dependencies, and sociocultural polarization as divisive forces in contemporary society. *Global Perspectives, 2*(1), Article 25658. https://doi.org/10.1525/gp.2021.25658

Strzelecki, M., Bayer, L., & Wlodarczak, A. (2025, January 3). Poland takes over EU presidency with focus on security. *Reuters*. https://www.reuters.com/world/europe/spats-divisions-mar-start-polands-eu-presidency-2025-01-03

The age gap in religion around the world. (2018, June 13). *Pew Research Center*. https://www.pewresearch.org/religion/2018/06/13/the-age-gap-in-religion-around-the-world

"This is a secular state": Polish government ends subsidies for religious publications. (2024, March 15). *Notes from Poland*. https://notesfrompoland.com/2024/03/15/this-is-a-secular-state-polish-government-ends-subsidies-for-religious-publications

Tomaszewski, M. (2021, March 9). Zamiast "osoba z macicą" powiedzieli "kobieta." Teraz Partia Razem kaja się i idzie na szkolenie. *Niezalezna*. https://niezalezna.pl/polska/zamiast-osoba-z-macica-powiedzieli-kobieta-teraz-partia-razem-kaja-sie-i-idzie-na-szkolenie/384404

Törnberg, P. (2022). How digital media drive affective polarization through partisan sorting. *Proceedings of the National Academy of Sciences, 119*(42), Article e2207159119. https://doi.org/10.1073/pnas.2207159119

Triandafyllidou, A., Meer, N., & Modood, T. (Eds.). (2011). *European multiculturalism: Cultural, religious and ethnic challenges*. Edinburgh University Press.

Tusk, D. (2025). *The speech of Polish Prime Minister at the European Parliament: "Europe is not yet lost as long as we are alive"* [Speech transcript]. Polish Presidency of the Council of the European Union. https://polish-presidency.consilium.europa.eu/en/news/the-speech-of-polish-prime-minister-at-the-european-parliament-europe-is-not-yet-lost-as-long-as-we-are-alive

Van Aelst, P., Strömbäck, J., Aalberg, T., Esser, F., de Vreese, C., Matthes, J., Hopmann, D., Salgado, S., Hubé, N., Stępińska, A., Papathanassopoulos, S., Berganza, R., Legnante, G., Reinemann, C., Sheafer, T., & Stanyer, J. (2017). Political communication in a high-choice media environment: A challenge for democracy? *Annals of the International Communication Association, 41*(1), 3–27. https://doi.org/10.1080/23808985.2017.1288551

Van Hauwaert, S. M., & Van Kessel, S. (2018). Beyond protest and discontent: A cross-national analysis of the effect of populist attitudes and issue positions on populist party support. *European Journal of Political Research, 57*(1), 68–92.

van Zomeren, M., Postmes, T., & Spears, R. (2008). Toward an integrative social identity model of collective action: A quantitative research synthesis of three socio-psychological perspectives. *Psychological Bulletin*, *134*(4), 504–535.

Verkuyten, M. (2013). *Identity and cultural diversity: What social psychology can teach us.* Routledge.

Verkuyten, M., & Martinovic, B. (2012). Immigrants' national identification: Meanings, determinants, and consequences. *Social Issues and Policy Review*, *6*(1), 82–112.

Viala-Gaudefroy, J. (2020). The evil savage other as enemy in modern US presidential discourse. *Angles*, *10*. https://doi.org/10.4000/angles.498

Wachter, M., Ventriglio, A., & Bhugra, D. (2015). Micro-identities, adjustment and stigma. *International Journal of Social Psychiatry*, *61*(5), 436–437.

West, E. A., & Iyengar, S. (2022). Partisanship as a social identity: Implications for polarization. *Political Behavior*, *44*(2), 807–838.

Wilson, A. E., Parker, V. A., & Feinberg, M. (2020). Polarization in the contemporary political and media landscape. *Current Opinion in Behavioral Sciences*, *34*, 223–228.

Ylä-Anttila, T. (2017). Familiarity as a tool of populism. *Acta Sociologica*, *60*(4), 342–357.

Young, D. G. (2023). *Wrong: How media, politics, and identity drive our appetite for misinformation.* Johns Hopkins University Press.

Żółciak, T., & Osiecki, G. (2022, December 21). Wyborcy opozycji chcą jednej listy. Na dziś wspólny start wydaje się jednak mało prawdopodobny. *Gazeta Prawna.* https://www.gazetaprawna.pl/wiadomosci/kraj/artykuly/8614830,wspolna-lista-opozycji-sondaz-wybory-do-sejmu.html

Zollinger, D. (2024). Cleavage identities in voters' own words: Harnessing open-ended survey responses. *American Journal of Political Science*, *68*(1), 139–159.

About the Authors

Piotr Kłodkowski is professor at the Centre for Comparative Studies of Civilisations at Jagiellonian University in Kraków. He has been a visiting professor at the University of Rochester, New York (2019, 2021, 2023, 2025) and is a former ambassador of Poland to India (2009–2014). He specializes in political science and international relations.

Małgorzata Kossowska is full professor in psychology and head of the Center for Social Cognitive Studies at Jagiellonian University. Her areas of interest include the cognitive and motivational underpinnings of complex social phenomena—e.g., political ideology, knowledge resistance, and the societal consequences of digitalization, particularly its impact on social cohesion and youth radicalization.

Anna Siewierska is professor at the Institute of Political Science, University of Rzeszów. As a political scientist, she focuses on the psychology and language of politics, political mythology, and the impact of new media, especially regarding radicalization and social polarization. She is the author of *The Language of Polish Politics* and *The Politics of Fear*.

cogitatio

ARTICLE

"Funny Weapons": The Norms of Humour in the Construction of Far-Right Political Polarisation

Gabriel Bayarri Toscano [1] **and Concepción Fernández-Villanueva** [2]

[1] Audiovisual Communication, Rey Juan Carlos University, Spain
[2] Social Psychology, Complutense University, Spain

Correspondence: Concepción Fernández-Villanueva (cfvillanueva@cps.ucm.es)

Abstract

Jair Bolsonaro in Brazil, Nayib Bukele in El Salvador, and Javier Milei in Argentina are just a few paradigmatic cases that represent, to different degrees, the rise of populism, the advances of right-wing radicalism, and the resurgence of extreme nationalism in Latin America in the last decade. The question that arose after the victory of the far-right was: How could this have happened? One of the instruments that undoubtedly contributed to this unexpected victory was a peculiar aspect of these political campaigns: memetic communication. Through the use of memes in social media (above all WhatsApp), the far-right transformed violent discourses against political opponents, feminism, racialised persons, and poverty into a series of messages legitimised through humour and irony. This process operated as a simplification that disrupted stable systems of social norms and metaphorical frameworks. Between September 2022 and February 2024, in the weeks leading up to and following each presidential election, we collected and analysed visual data employing open-source software. We also conducted ethnographic fieldwork and digital ethnography during the weeks preceding the elections to capture online and offline discourses and the affective milieu of each electoral campaign, providing contextual insight into the impact of memetic communication. Our analysis demonstrates the trivialisation and legitimisation of violence against political opponents and other social groups. This process may also be interpreted as an attempt to render the unconstitutional as legitimate, framing exclusionary or violent political acts as necessary or even virtuous. Much of this legitimisation was camouflaged under the mask of supposed humour and irony, which in reality was insulting, prejudicial, and dehumanising.

Keywords

far-right; humour; Latin America; memetic communication; social norms; violence

1. Introduction

It was 2 PM in a suburb of Rio de Janeiro in October 2022. A skinny, ownerless dog was begging for food at a demonstration in support of Bolsonaro. A protester shouted: "Watch out for that dog—if Lula wins, we'll have to call him picanha!"—referring to a prized cut of beef. Around him, dozens of people erupted in laughter, amplifying the joke: "If Lula wins, he'll be the first to eat it!" shouted a man dressed in yellow. "We'll have to be careful when buying a hot dog!" added a young man wearing a T-shirt with their leader's face. As Bolsonaro appeared on stage, thousands of Brazilian flags—green and yellow—were raised to the cry of "Mito! Mito! Mito!" erasing any trace of the street dog from their minds.

Elsewhere in Latin America, in a suburb of Buenos Aires, a supporter of Milei during his campaign in September 2023 remarked: "Milei is good, he is a man who likes dogs; his dogs are part of his thinking." The boy, attending the march with his Uber Eats bike, was joined by another voice: "But the Kirchneristas eat cats; soon we'll be eating dogs for sure if that mafia continues to rule," exclaimed a middle-aged man, visibly anxious. "We run the risk of becoming like Venezuela—a country of laughter, a communist country," an elderly woman remarked, oscillating between laughter and indignation. These unsettling jokes travelled beyond Latin America, taking on new meanings. By 2024, Donald Trump, in a presidential debate against Kamala Harris, claimed that Haitians living in the US "are eating dogs." Harris's initial reaction was laughter—an instinctive, perhaps nervous, response to what she perceived as rhetorical absurdity. However, her expression quickly changed as she attempted to reframe the comment within the domain of the "serious."

The trajectory of these jokes illustrates how humour can serve as a social norm—either to strengthen in-group bonds and alleviate anxiety or, conversely, to justify the violation of norms. Within the far-right political ecosystem, humour operates not simply as entertainment but as a strategic tool to reconfigure the boundaries of what is socially acceptable. Leaders such as Bukele, Milei, Bolsonaro, and Kast are not merely aligned in their militarism, authoritarian tendencies, violent rhetoric, and market fundamentalism; they are also bound by their deployment of humour as a means of shaping political discourse. As influencers in their own right, they cultivate and amplify humorous narratives on social media, where laughter becomes a mechanism for normalising otherwise transgressive discourses.

In this article, we analyse how memes produced by supporters of far-right leaders—particularly in Brazil, Argentina, and El Salvador—contribute to the normalisation of violence and the construction of exclusionary social norms. Through a contextualised visual analysis of meme communication, we examine the interplay between humour and political discourse, with a focus on how visual and affective framings legitimise aggression, trivialise violence, and reshape collective perceptions of what is socially acceptable. By exploring both the discursive and emotional dimensions of these memes—whether manually created or AI-generated—we argue that humorous political content functions as a powerful tool in the radicalisation of right-wing movements and the reinforcement of authoritarian narratives within the Latin American postcolonial context.

2. The Normalisation of Violence Through Norms of Humour

2.1. The "Serious" vs. the "Humorous"

Discursive coding in terms of "seriousness" or "humour" has traditionally been framed as a dichotomy, separating direct, literal communication from humorous, interpretive communication (Wodak et al., 2021). Humour, as a communicative code, is distinguished by its flexibility—it allows for multiple interpretations and the suspension of credibility in the transmission of messages (Dynel & Messerli, 2020). However, this very adaptability makes humour a powerful mechanism for shaping social norms, influencing how audiences perceive what is acceptable within political discourse (Shifman, 2019).

One of the most concerning aspects of humour in political contexts is its capacity to mask misinformation and desensitise audiences to forms of symbolic violence (Haslam, 2022). The claim that political opponents or marginalised groups would resort to eating dogs, for instance, is a rhetorical device that simultaneously ridicules and dehumanises (Billig, 2019). In Brazil, some demonstrators who laughed at such jokes later claimed they did not actually endorse such an idea, while others insisted it was simply a means of mocking political adversaries. Regardless of intent, these jokes serve to reinforce a normative framework in which violence and exclusion are rendered trivial or even humorous. This aligns with Cardoso de Oliveira's (2004) notion of "dignified moral substance"—humour here operates to strip political opponents of their moral standing, justifying their exclusion from the realm of legitimate politics by portraying them as fundamentally transgressive (Fielitz & Ahmed, 2021).

2.2. Violence and Humour

The use of humour to normalise violence has become a core strategy within far-right political communication (Askanius & Keller, 2021). Whether through grassroots digital spaces or official party messaging, humour allows for the introduction of otherwise unacceptable discourses into the public sphere (Merrill et al., 2024). It facilitates discussions of topics such as capital punishment, misogyny, homophobia, and xenophobia by presenting them in an ironic or exaggerated manner (Burkart et al., 2023). This process gradually shifts social norms, making these positions appear more acceptable over time (Brantner et al., 2019).

Humour also provides a convenient shield for those who propagate violent ideas. When challenged, individuals can retreat behind the argument that it was "just a joke," thus evading moral and political accountability (Billig, 2019). This defensive mechanism creates a culture in which those who critique violent humour are themselves ridiculed as overly sensitive or lacking in social awareness (Parrot & Hopp, 2020). As a result, humour functions as a gatekeeping tool, delineating who belongs within certain ideological communities and who does not (Gallardo-Paúls, 2018).

Far from being a passive cultural phenomenon, humour actively shapes the political agenda (Cho et al., 2020). Through repeated exposure to dehumanising jokes, audiences become more receptive to explicit violence (Bastian & Haslam, 2011). First, humorous framings test the boundaries of what is acceptable; later, explicit policy positions or violent rhetoric are introduced with reduced resistance (Pérez & Greene, 2016). This process underscores humour's role as a transitional mechanism—an intermediary step between symbolic violence and its material enactment (Denisova, 2019).

2.3. Latin American Particularities and the Postcolonial Context

Understanding humour as a vehicle for normalising implicit violence requires situating it within its cultural and historical context (Pinheiro-Machado & Vargas-Maia, 2023). In Latin America, the postcolonial condition is key to deciphering how humour operates within political polarisation. Jokes that suggest "a good bandit is a dead bandit" (Bolsonaro's Brazil), that "a dictator can be cool" (Bukele's El Salvador), or that "selling organs is just another form of market freedom" (Milei's Argentina) must be analysed in relation to the region's history of authoritarianism, structural inequality, and racial hierarchies (Bassil et al., 2023).

The militarisation of public discourse is particularly relevant here (Fernández-Villanueva & Bayarri, 2021a). In Latin American far-right humour, opponents are frequently depicted as animals, criminals, or existential threats to the nation (Fernández-Villanueva & Bayarri, 2021b). This tactic—rooted in historical practices of dehumanisation—renders political enemies disposable. The friend-enemy distinction is reinforced through popular memes and jokes that depict progressive leaders as rats, cockroaches, or parasites. Such imagery is not new; it echoes fascist and Nazi propaganda of the 20th century, which similarly employed humour to prime audiences for exclusionary violence (Fielitz & Thurston, 2018).

Humour, then, is far from innocuous—it is a communicative code that both reflects and constructs social norms (Shifman, 2019). In fragile democratic contexts, where reactionary movements wield increasing influence, humorous discourses can become key instruments of radicalisation (Askanius, 2021). The far-right's strategic use of humour in Latin America demands urgent scholarly attention, particularly regarding regulation and digital literacy (Gomes-Franco e Silva et al., 2022). The normalisation of "soft" codes of exclusion through humour not only erodes democratic debate but also sets the stage for more explicit forms of political violence (Colussi et al., 2023).

3. Memetic Communication and the Far-Right

3.1. From Political Cartoons to Digital Memes: Historical Shifts in Visual Political Commentary

In recent years, political discourse in Latin America has become increasingly polarised, with electoral periods serving as moments of peak rhetorical confrontation (Merrill et al., 2024). Political polarisation seeks to legitimise violence against opponents by classifying them as enemies, rendering them politically and socially expendable (Slimovich, 2022). Bandura (2019) highlights how this process fosters devaluation and dehumanisation, elevating certain leaders to near-mythical status while demonising their adversaries— a dynamic that involves the sanctification of political leadership and the simultaneous degradation of dissenting groups, including non-political and marginalised identities. As Soler Gallo (2019) argues, the division of society into an "us" versus "them" dynamic is reinforced through fear, manipulation, hyperbolic statements, irony, sarcasm directed at opponents, and the delegitimisation of conventional politics. This phenomenon has been exacerbated by digital communication technologies, which amplify political narratives through new forms of media (Rubira & Puebla, 2018).

Among these new forms, memetic communication has emerged as a particularly influential mode of discourse, deeply intertwined with the evolution of social norms. However, the memetic mode represents a break from previous forms of political visual commentary, such as printed political cartoons. While both genres deploy

humour and visual exaggeration, traditional editorial cartoons—typically authored, signed, and published in institutional media—often framed their satire within the bounds of journalistic commentary, civic pedagogy, or political critique. Their reach was largely controlled by editorial standards and limited by circulation.

Memes, in contrast, are decontextualised, anonymous, and virally distributed, unanchored from institutional oversight. They rely less on caricatured drawing styles and more on the remix of photographs, screenshots, and AI-generated images. Their humour is not always oriented toward satire, but toward affective mobilisation, mockery, and symbolic annihilation. While the political cartoon is episodic and author-driven, the meme is accumulative, networked, and participatory. It thrives on immediacy and replication, allowing users to appropriate, distort, or intensify its meaning. As noted by Shifman (2019), the hyper-signification of memes allows for layered interpretations that transcend the linear communication style of traditional cartoons. Furthermore, their multimodal design (Burkart et al., 2023; Hakoköngas et al., 2020) enables them to engage audiences simultaneously on cognitive, visual, and emotional levels, making them especially potent tools for shaping social perception in volatile political climates.

3.2. Humour, Norms, and Polarisation in Contemporary Far-Right Memetic Discourse

Memetic communication does not merely reflect political divisions; it actively constructs and reinforces social norms. Unlike traditional political communication, which relies heavily on rational discourse and structured argumentation, memes operate primarily through metaphor, metonymy, and affective appeals. They often use emotionally charged imagery or culturally loaded symbols, and their brevity enhances their ability to spread virally. Their effects are diffuse but powerful, resonating with audiences on an emotional rather than rational level. Moreover, memes can be spontaneously shared and remixed, connecting with existing content in ways that are difficult to predict or control (Bredekamp, 2017). This emotional immediacy makes memes particularly effective tools for fuelling polarisation and reinforcing ideological narratives. Digital humour plays a central role in defining what is socially acceptable. By repeatedly circulating memes that ridicule, dehumanise, or delegitimise political opponents, these visual artefacts contribute to the gradual erosion of democratic norms, normalising hostility, aggression, and political violence.

The study of memetic communication is relatively recent, but it is increasingly recognised for its role in shaping political narratives (Arkenbout, 2022; Askanius & Keller, 2021). According to Dynel and Messerli (2020), the most common form of internet meme consists of an image paired with text, designed to elicit a humorous response. This humour is multimodal (Burkart et al., 2023; Hakoköngas et al., 2020), engaging viewers cognitively, visually, and emotionally. Shifman (2019) notes the phenomenon of hyper-signification, in which memes layer meanings and condense complex social critiques into simple formats. These compressed, emotionally loaded messages act as powerful mobilising forces, attaching feelings to digital objects and contributing to the renegotiation of shared social norms (Cho et al., 2020).

A key aspect of this process is the way humour disguises violence and offence as entertainment. Political memes frequently cross the line into hate speech, but their comedic framing offers plausible deniability. Individuals can evade accountability by claiming they were "just joking," while critics are dismissed as humourless or overly sensitive (Billig, 2019). This tactic is particularly potent in jurisdictions where legal definitions of hate speech remain narrow or ambiguous (Figueruelo Borrieza & Martín Guardado, 2023). In the Latin American context, humour has long served as a means of reinforcing social hierarchies,

especially those based on race, gender, and class. Memes that depict opponents as animals, degenerates, or existential threats function not only as ideological expressions, but as performative acts that delineate who belongs in the national community—and who does not.

In Brazil, Argentina, and El Salvador, memetic communication has become integral to far-right political discourse. In Brazil, Bolsonaro's online supporters frequently posted memes that portray adversaries as weak, corrupt, or effeminate, reinforcing traditional gender norms and delegitimising dissent. In Argentina, supporters of Javier Milei have used memes to frame critics as parasites, equating economic reform with moral purification. In El Salvador, Bukele's online presence features memes that depict him as a saviour or messiah, while opposition figures are mocked as traitors or puppets of foreign powers.

A particularly salient rhetorical device is exaggeration, which involves amplifying specific traits or events beyond their realistic proportions to provoke strong emotional responses or ridicule. This often manifests as hyperbole, a deliberate and overt overstatement used to intensify the message's impact and elicit humour or indignation (Wodak et al., 2021). While all hyperboles are exaggerations, not all exaggerations rise to the level of hyperbole, which is typically more stylised and explicit in its distortion. In the context of far-right memes, these techniques distort opponents' traits for comic effect or vilification. Although satire can sometimes serve as critique, in this case, it becomes a vehicle for rhetorical violence (Gallardo-Paúls, 2018). Memetic mockery has real-world implications: It can harm reputations, incite harassment, or serve as justification for discriminatory policies (Pérez & Greene, 2016). Moreover, the alignment of meme culture with populist rhetoric further amplifies its power. Populist discourse thrives on simplified narratives of good versus evil, friend versus enemy (Mouffe, 2022), and memes provide the perfect medium for transmitting such binaries.

On a global scale, populist movements deploy visual simplifications that echo 20th-century propaganda techniques—especially those used by fascist and communist regimes (Bassil et al., 2023). These movements use humour to present exclusionary ideologies as common sense, masking radicalism with laughter. Leaders like Bolsonaro, Milei, and Bukele embrace this aesthetic as part of their public personas, turning memes into instruments of personal mythmaking and ideological mobilisation (Bayarri et al., 2024; Colussi et al., 2023). There is growing concern over how to regulate memetic communication in this context. As humour operates in legal and cultural grey zones, tech platforms and policymakers struggle to moderate meme-based disinformation and hate. Initiatives like "platform governance triangles" (Gorwa, 2019) involving governments, tech firms, and civil society have attempted to address this gap, but enforcement remains inconsistent. National laws often lag behind the digital innovation and cultural ambiguity of memes, allowing exclusionary or violent messages to circulate unchecked.

In the context of electoral politics in Latin America, memes have been instrumental in legitimising political violence—not through explicit calls to action, but by defining the terms of acceptable political discourse. Opponents are framed as not only wrong but subhuman, immoral, or laughable. This affects not only the targets of such representations—who may be women, LGBTQI+ people, indigenous communities, or the poor—but also the broader public understanding of who belongs in the political conversation (Denisova, 2019; Fraticelli, 2023).

Academic studies (e.g., Bowen, 2020; Merrill & Lindgren, 2021) confirm that memes are no longer peripheral to political communication. They play a central role in shaping public opinion and in producing cultural

norms. In Latin America, the stakes of this memetic war are particularly high. In Brazil, memes contributed to Bolsonaro's image as a protector of the family and nation. In Argentina, they helped cast Milei's opponents as enemies of progress. In El Salvador, they presented Bukele as a messianic leader whose authority cannot be questioned. These are not just jokes—they are "funny weapons" that change the rules of political engagement.

4. Objectives and Methods

In this article we aim to analyse how memes produced by supporters of far-right leaders—particularly in Brazil, Argentina, and El Salvador—contribute to the normalisation of violence and the construction of exclusionary social norms. Specifically, we examine the interplay between humour and political discourse, focusing on the ways in which visual communication legitimises aggression, reinforces hierarchies, and reshapes perceptions of what is socially acceptable.

We also seek to examine how memetic representations of political antagonists contribute to the normalisation of violence and the transformation of social norms. Building on existing research on memetic communication, we explore how humour operates as a mechanism for legitimising exclusionary discourses, reinforcing political hierarchies, and shaping collective perceptions of what is socially acceptable. Furthermore, we investigate the role of generative AI tools—such as Stable Diffusion, MidJourney, DALL-E, and GANs—in reshaping the production and dissemination of political memes, assessing their impact on the intensification of political polarisation, the reinforcement of stereotypes, and the legitimisation of violence. This process may also be interpreted as an attempt to render the unconstitutional as legitimate, framing exclusionary or violent political acts as necessary or even virtuous, particularly against marginalised social groups.

The methodology integrates visual data collection with digital ethnography, drawing from Alonso's (1994) socio-hermeneutic approach, which advocates for contextual analysis of visual narratives beyond their explicit content, and Pauwels (2015), who highlights the role of images in generating new theoretical insights into social life. Visual discourse is seen as a social practice embedded with the intentions and positions of its actors, revealing ideological structures (Barthes, 1995; Serrano & Zurdo, 2023). Data collection focused on open Telegram groups where far-right supporters actively engage in political discourse. The collection took place in the weeks preceding and following each presidential election in Brazil, Argentina, and El Salvador, between September 2022 and February 2024. Specifically, six Telegram groups were analysed, with two per country: one associated with the "official account" of a right-wing populist politician and another representing a self-managed grassroots group.

These spaces provide insight into the memetic strategies employed by supporters of Bolsonaro in Brazil, Milei in Argentina, and Bukele in El Salvador, offering a lens into how humour functions as a vehicle for reinforcing social norms of exclusion that extend beyond political opposition and target marginalised groups such as LGBTQI+ individuals or racial minorities, thereby fostering everyday or interpersonal polarisation and legitimising forms of political violence. In Brazil, data collection covered the electoral rounds held on October 2 and October 30, 2022. The "Carlos Bolsonaro—Official Account" group, with 144,267 members, contributed 72 images, while the "Bolsonarista" group, with 38,362 members, provided 17,072 images. For Argentina's elections on October 22 and November 19, 2023, data collection spanned from October 1 to November 30. The "Javier Milei Official" group (675 members) contributed 1,564 images, while the "Milei

president 23–27. Long Live Freedom" group (481 members) contributed 6,202 images. In El Salvador, data was gathered between December 4, 2023, and February 4, 2024, coinciding with the single-round election on February 4. The "Nuevas Ideas" group (140 members) contributed 118 images, while the "Nayib Bukele" group (3,333 members) provided 518 images.

We applied critical visual methodology, adapting grounded theory to the visual domain, as proposed by Mey and Dietrich (2016). The process followed these steps: (a) contextualisation of data aligned with the research questions; (b) creation of a sample inventory from the Telegram datasets; (c) manual categorisation of recurring themes; (d) image coding; and (e) interpretive theorisation of the memes in relation to political polarisation. We selected images that exemplify dominant narrative structures rather than attempting an exhaustive review of all collected data. The categories were cross-validated by both researchers to ensure analytical consistency.

To complement the visual data, ethnographic and digital ethnographic fieldwork was conducted to capture both online and offline discourses, as well as the affective climate surrounding each election. Offline data were collected through participant observation at political rallies and informal interviews with attendees. These interactions offered insights into how online humour is rearticulated in public discourse. Memes encountered in Telegram groups were frequently referenced in chants, signs, and casual conversation, confirming the circulation of memetic messages across communicative contexts. This contextual approach enhances the understanding of how memetic communication functions as a mechanism for consolidating political identities, reinforcing social hierarchies, and legitimising political aggression. By situating memes within broader cultural and political frameworks, the study offers a critical perspective on their role in shaping contemporary social norms and the boundaries of political acceptability.

5. Results. The Role of Social Norms in Memetic Communication: Brazil, Argentina, and El Salvador

The analysis of meme-based communication in Argentina, Brazil, and El Salvador reveals that generative AI memes tend to be more generic and symbolic, often lacking textual content and expressing broader, more universally recognisable symbols. Manual memes, by contrast, are more context-specific, reactive to immediate political events, and often carry more emotionally charged, even sadistic content. In our sample, AI-generated images represented between 0.5% and 6% of the visual material, while manual memes accounted for the overwhelming majority, ranging from 94% to nearly 99% depending on the group.

These differences suggest that while both types reinforce similar social norms related to exclusion and violence, AI-generated memes may be more prone to polarisation via exalted imagery of leaders and dehumanising portrayals of opponents (Bayarri & Fernández-Villanueva, 2025). Thematically, two motifs recur with high frequency: the sanctification of political leaders, particularly in AI-generated content, and the dehumanisation of adversaries, more prominent in manually crafted memes. These patterns appeared consistently across the three countries, with religious or mythological imagery, animalistic portrayals, and calls for exclusion or mockery.

Audience engagement metrics within the Telegram groups—measured by reactions, replies, and shares—indicate that memes that use humour to delegitimise opponents tend to attract higher interaction. Popular memes were often adapted by users, shared across other platforms, or remixed with new captions or filters,

suggesting a participatory dynamic that reinforces the normative power of humour. In several instances, memes that gained traction online were later cited or echoed during offline rallies, indicating an intertextual feedback loop between digital and analog spaces.

Ethnographic fieldnotes from rallies and political gatherings reinforce this connection. Supporters frequently referenced memes during interviews or shouted popular slogans derived from meme culture. This offline uptake demonstrates how digital humour not only shapes online discourse but also informs real-world political identity and allegiance. By studying both the circulation and public uptake of memes, the findings show how memetic humour contributes to norm formation by testing, validating, and reinforcing social boundaries through collective laughter.

5.1. Argentina: Nationalism, Heroism, and the Normalisation of Political Exclusion

In Argentina, both generative AI and manually created memes reinforce a nationalist and populist imagery centred on Javier Milei and his ideological narratives. In the "Javier Milei Official" and "Milei President 23–27. Long Live Freedom" Telegram groups, AI-generated images constitute 4.86% and 3.53% of the content, respectively. These AI-generated visuals depict Milei as a heroic, lion-like figure, invoking themes of medieval heroism, patriotism, and anti-establishment defiance—representing a constructed norm in which leadership is equated with masculine strength and aggressive individualism (see Figure 1). For instance, the image on the left shows the "lion Milei" acting fiercely against his political opponents, who are dehumanised and transformed into rats, while the central image portrays Milei as a Napoleonic figure. In contrast, the image on the right, which was produced manually, depicts Peronist Sergio Massa talking to a sympathiser—also animalised—to whom he says, "Do you understand that I devalue the coin so that the right does not come back?" to which the sympathiser answers, "Yes, long live Perón."

Figure 1. Representations of Milei's heroism and attacks on Peronism. Source: Own composition.

Manually created memes (95.14% and 96.47%) take a more combative approach, often ridiculing political opponents and reinforcing the notion that traditional media and left-wing movements threaten national integrity. The humour embedded in these memes serves not only as entertainment but as a form of ideological reinforcement, establishing mockery as a normative means of delegitimising adversaries. In our corpus, over 60% of the memes included explicit derision of political opponents, with common labels such as "rats," "leeches," or "scum."

Responses in Telegram groups reflect this process: Supporters praise Milei as a "Templar" and "national hero," while derogatory remarks about opponents reinforce social norms of exclusion that extend beyond political opposition and target marginalised groups such as LGBTQ+ individuals or racial minorities, thereby fostering everyday or interpersonal polarisation.

These same narratives are echoed in real-world political discourse. At Milei's rallies, humour and aggression intersect in meme discourse that is often charged with emotions of rejection, anger, and hostility. Supporters reassert the memes' messages in offline settings. "Milei is a beast, a lion who will save us," a supporter states, connecting the digital meme discourse with political mobilisation. By repeatedly circulating memes that establish humour as a socially acceptable medium for exclusion, these spaces redefine what is legitimate political speech and what is considered too "serious" to challenge.

5.2. Brazil: Religious Morality and the Sanctification of Political Leadership

Brazilian far-right memetic communication relies heavily on religious and nationalistic themes to reinforce conservative social norms. Within the "Carlos Bolsonaro—Official Account" and "Bolsonaristas" Telegram groups, the use of generative AI in meme production is significantly lower—2.78% and 0.51%—indicating a greater reliance on traditional meme formats. Despite this, manually created memes are highly effective in shaping discourse, reinforcing notions of Bolsonaro as Brazil's moral guardian.

Humour plays a crucial role in constructing norms around political legitimacy. The juxtaposition of innocence and corruption in Bolsonaro-related memes reinforces a binary opposition of "good" versus "evil," normalising the idea that his opponents are inherently corrupt or morally depraved. AI-generated imagery amplifies this message by portraying Bolsonaro as a sanctified leader, aligned with a divine vision of omnipotent strength and moral absolutism (see Figure 2). The two images on the left exemplify this aesthetic, presenting Bolsonaro as a divinised and enlightened figure, evoking Christian-nationalist connotations of sacred leadership. In contrast, the manually produced image on the right targets the LGTBIQ+ community through grotesque humour: It depicts a trans woman who, instead of engendering a baby, would be engendering an excrement—yet another display of exclusionary nationalism through ridicule.

Figure 2. Representations of Bolsonaro's divinity and attacks on LGTBIQ+ communities. Source: Own composition.

Reactions within these groups indicate strong affective ties to these representations: The most interacted memes included either religious blessings, mythological symbolism, or violent commentary against opposition

figures. Over 70% of the memes analysed used some form of religious or moral framing. Bolsonaro's followers frequently express solidarity using religious language: "All together with the patriot," "Myth, let's not forget him." Conversely, reactions to memes depicting opponents often involve calls for violence—"That vagrant should have been killed," "Shot in the head"—illustrating how humour, when integrated into political discourse, helps to redefine violent rhetoric as an acceptable norm.

Offline, these messages resonate strongly. During campaign events, supporters echoed meme slogans, chanted biblical references, or wore imagery identical to popular digital memes. "I support Bolsonaro because he supports the family," a voter declares, highlighting how digital humour fosters ideological cohesion and frames political loyalty in moral terms. This fusion of meme culture and religious nationalism reflects the broader transformation of far-right discourse into an emotionally charged, exclusionary system.

5.3. El Salvador: AI-Driven Hero Worship and the Rewriting of Political Legitimacy

El Salvador's memetic discourse differs in its higher rate of generative AI use (6.16%, the highest among the three countries), demonstrating a sophisticated visual strategy for constructing Nayib Bukele's image as a near-divine leader. The "Nuevas Ideas" and "Nayib Bukele" Telegram groups prominently feature AI-generated portrayals of Bukele as the embodiment of national salvation, reinforcing a normative vision of leadership that merges technological progress with authoritarian control—an aesthetic that sacralises the figure of the leader as a technological deity (see Figure 3). The two images on the left reflect this approach: the first attributes superpowers to President Bukele, while the second celebrates an alleged moment of national prosperity under his rule. In contrast, the manually produced image on the right shows Bukele physically assaulting a "gang member" while a group of journalists writes critically in the background—suggesting that the media ignores the supposed security benefits of Bukele's actions. The caption reads "and then they hit him with everything," implying that journalists only focus on condemning violence without acknowledging its effects.
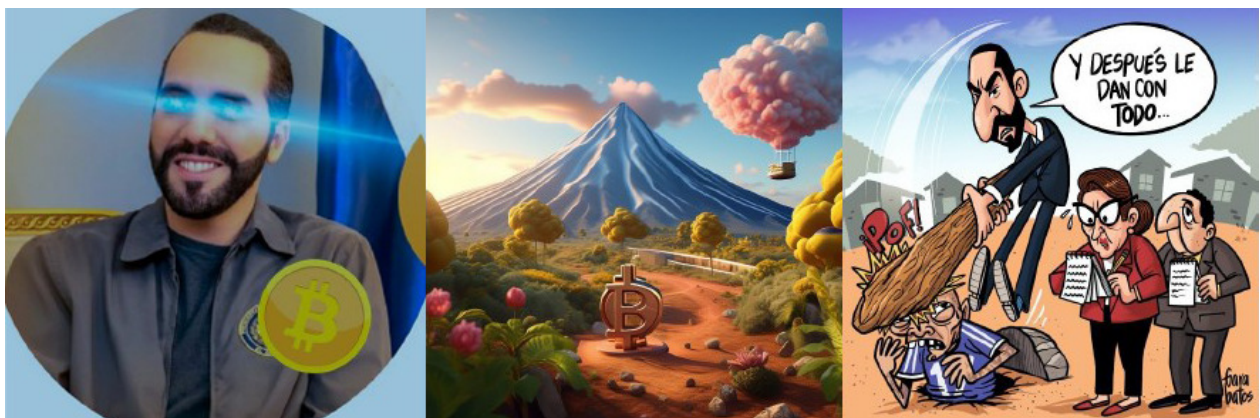
Figure 3. Representation of El Salvador under Bukele. Source: Own composition.

Manual memes, by contrast, aggressively target political opposition, frequently associating Bukele's adversaries with crime, economic ruin, or external threats. Over half of these memes reference opposition leaders as "puppets," "communists," or "rats." The frequent depiction of opposition figures as foreign agents or obstacles to national progress aligns with broader regional trends, where humour serves as a means of social control, reinforcing norms that exclude critics from legitimate political participation. Audience

engagement in these groups shows strong enthusiasm for heroic portrayals of Bukele. "The best president in Latin America. Of the world," one comment reads. "He has saved us all. He is Jesus," another asserts. These statements illustrate how humour and AI-generated aesthetics contribute to the sacralisation of Bukele's leadership, entrenching social norms that position him as the sole figure capable of ensuring national prosperity.

Offline, these norms persist in public sentiment, where Bukele's supporters adopt memes as part of everyday discourse. "The coolest president," one campaign volunteer states, while another compares El Salvador's current security situation to its violent past, legitimising Bukele's authoritarian measures as necessary corrections to historical disorder. These interactions reveal how digital culture informs not only voter identity but also popular imaginaries of leadership, national salvation, and political belonging.

6. Discussion and Conclusions: The Normalisation of Violence Through Humour in Memetic Communication

6.1. Humour as a Mechanism for Reinforcing Social Norms and Justifying Violence

Memetic communication in Brazil, Argentina, and El Salvador plays a pivotal role in the reconfiguration of social norms, particularly in the normalisation of political violence through humour. By repeatedly using satire, irony, and ridicule to frame political opponents as ideological threats, memes function as a rhetorical tool that transitions veiled violence into explicit aggression (Merrill et al., 2024). These dynamics establish humour not merely as a discursive strategy but as a means of defining what is socially permissible in political dialogue (Billig, 2019; Shifman, 2019).

Across the three countries, memes shape and enforce social norms by clearly opposing exalted images of the "us" group to degraded, dehumanised depictions of the "other," whether through animalisation, robotisation, or symbolic trashing. These framings, often linked to violence or exclusion, mirror historical methods of marginalisation and are deployed through ridicule and sarcasm, forming acts of everyday polarisation. Haslam (2022) argues that even subtle forms of interpersonal coldness can be socially damaging, while studies on humanisation (Bastian et al., 2013; Pavetich & Stathi, 2020) suggest that emotional connection mitigates these effects. Our findings support this view and link meme-based dehumanisation with broader transformations in everyday relationships between citizens. This is in line with Reynares' (2024) theory of everyday polarisation, where not only direct violence but also emotional coldness, mockery, and fear contribute to antisocial, depersonalising attitudes (see Figures 1, 2, and 3).

In Argentina, digital humour constructs political adversaries as existential threats to national purity, reinforcing a norm in which ideological dissent is framed as an attack on the nation itself (Bassil et al., 2023). In Brazil, memes integrate Christian morality into political discourse, presenting leftist opposition figures as fundamentally corrupt and morally bankrupt (Rubira & Puebla, 2018). In El Salvador, AI-generated memes elevate Bukele to an almost divine status, reinforcing a norm in which authoritarian governance is equated with stability and progress (Pinheiro-Machado & Vargas-Maia, 2023).

Far from being harmless expressions of online culture, memes serve as cultural artifacts that regulate social norms by embedding exclusion, mockery, and delegitimisation into everyday discourse. The humour they

employ lowers the threshold for accepting aggression, making political violence seem justifiable, if not inevitable (Fielitz & Ahmed, 2021). When opposition figures are continuously portrayed as deviant, inhuman, or dangerous, the transition from symbolic violence to material aggression becomes easier to rationalise (Denisova, 2019). This phenomenon illustrates how humour operates as a first step in shifting social acceptance toward more explicit forms of political violence (Gallardo-Paúls, 2018).

6.2. AI-Generated Memes and the Acceleration of Polarisation

The rise of generative AI in meme production further complicates this process. While AI tools such as Stable Diffusion, MidJourney, and DALL-E have democratised political meme creation, they have also intensified the speed and scale at which social norms can be shaped and manipulated (Bayarri & Fernández-Villanueva, 2025). The accessibility of these tools allows for the rapid production of visually sophisticated content, making political messaging more persuasive and immersive (Burkart et al., 2023). However, this technological shift also presents challenges in distinguishing authentic discourse from algorithmically generated propaganda (Brantner et al., 2019).

Our dataset revealed significant differences in the formal and rhetorical features of AI-generated and manually created memes. AI memes tend to feature fewer words and rely more on symbolism, emphasising epic or mythologised portrayals of political figures. These images often depict leaders as divine, animalistic, or superhuman—particularly lions, messianic beings, or armored warriors—and are often rendered with high aesthetic quality. Such depictions contribute to the aspirational emotional appeal of these figures and serve to elevate their perceived legitimacy beyond normal political critique.

In contrast, manually created memes are frequently reactionary, rougher in visual execution, and replete with textual elements such as insults, slang, and explicit polarising cues. They focus on degrading opponents through ridicule, distortion, or associations with criminality, moral decay, or foreign influence. These memes are often more immediately linked to current events and adapted in real time, allowing users to respond quickly to breaking news and share content that resonates with shared grievances or frustrations.

Interestingly, while AI-generated memes constituted a smaller portion of the dataset (between 2.78% and 6.16%, depending on the country), user interactions with them—measured through the number of reactions, reposts, or positive emoji responses—suggested higher affective reinforcement. They were less frequently modified by users but more frequently commented on as "powerful," "epic," or "moving." In contrast, manually created memes were more likely to be remixed and reshared with personalised captions or annotations, demonstrating a higher degree of participatory engagement and narrative adaptability.

These findings suggest that AI-generated memes are particularly effective at conveying hegemonic and aspirational imagery, reinforcing long-term symbolic associations between authoritarian leadership and national salvation. Meanwhile, manually created memes remain central to real-time political antagonism and the construction of in-group/out-group boundaries. Together, both forms of meme production contribute to the affective and discursive terrain of polarisation—albeit through different strategies and temporal dynamics.

Although AI enhances engagement with political content, it also exacerbates the oversimplification of political discourse. When memes reduce complex political debates to emotionally charged images, they

solidify social norms that favour simplistic, binary understandings of governance—where opposition is framed as inherently corrupt or dangerous, and where leadership is elevated to near-mythical status. This form of digital communication reinforces political divisions, making it more difficult for democratic dialogue to thrive (Cho et al., 2020).

6.3. The Challenge of Regulating AI-Driven Memetic Communication

The widespread adoption of AI in meme production presents significant regulatory and ethical challenges. The ability to generate highly persuasive political content at scale increases the risk of misinformation and the strategic manipulation of public opinion (Gomes-Franco e Silva et al., 2022). Given Latin America's history of political instability and economic inequality, AI-generated memes tend to widen existing digital literacy gaps by privileging those with greater access to digital tools and interpretive skills, while rendering other populations—particularly older adults, rural communities, and individuals with lower levels of formal education—more susceptible to these persuasive visual narratives. As Fielitz and Thurston (2018) argue, such memes operate through aesthetic cues and affective triggers that bypass analytical reasoning, making it harder for less digitally literate audiences to identify bias, satire, or manipulation.

In contexts where educational inequality, limited media literacy, and high levels of digital consumption converge—as in many parts of Latin America—certain groups are more likely to interpret memetic messages literally or as truthful representations, particularly when they mirror cultural narratives already present in everyday discourse. This makes AI-driven visual propaganda especially potent among populations with limited access to critical media education.

In authoritarian-leaning democracies, where trust in traditional media is already compromised, AI-generated memes can serve as an alternative propaganda tool, allowing leaders to bypass institutional checks and consolidate their image through visual storytelling. This is particularly evident in El Salvador, where Bukele's memetic presence is heavily AI-enhanced, positioning him as a transformative figure beyond traditional political constraints (Mouffe, 2022). In Brazil and Argentina, AI-driven meme culture similarly serves to reinforce ideological narratives, deepening political antagonisms and fostering segmented information bubbles (Parrot & Hopp, 2020).

Given these developments, there is an urgent need for a multifaceted approach to mitigate the risks posed by AI-driven memetic communication. Strengthening media literacy programs is crucial to equipping citizens with the ability to critically assess digital content (Haslam, 2022). Additionally, regulatory frameworks must be adapted to address the potential for AI-generated disinformation, ensuring transparency in the production and dissemination of political memes. While outright censorship may not be a viable solution, fostering digital accountability through platform governance and content moderation remains a critical step in mitigating the adverse effects of AI-driven propaganda (Wodak et al., 2021).

6.4. The Role of Memes in Shaping Future Social Norms

Memetic communication has become a fundamental mechanism in defining the contours of political legitimacy in Brazil, Argentina, and El Salvador. As memes continuously shape perceptions of leadership, opposition, and ideological affiliation, they contribute to the gradual acceptance of aggressive political rhetoric (Kien, 2019).

Humour, in this context, is not neutral—it serves as a powerful instrument that establishes what forms of discourse are permissible, influencing public attitudes toward political violence (Parker, 2019).

As Mason (2018) notes, polarisation impairs democratic negotiation by fostering emotional hostility rather than ideological difference. Haslam (2022) also warns that emotional detachment and interpersonal coldness are among the damaging outcomes of everyday polarisation. AI-generated memes amplify these effects, accelerating the normalisation of exclusionary politics and reinforcing the perception that aggression against certain groups is not only acceptable but necessary. This process highlights the urgent need to understand memes not merely as digital ephemera but as core instruments in the transformation of social norms (Fraticelli, 2023).

As technology continues to evolve, the intersection of political humour, AI, and memetic communication will remain a crucial area of study. Addressing these challenges requires a concerted effort to balance digital freedom with ethical responsibility, ensuring that memetic spaces do not become unchecked arenas for the legitimisation of political violence (Gorwa, 2019). Ultimately, fostering a more informed and pluralistic media environment will be key to resisting the entrenchment of exclusionary and aggressive social norms in digital political culture.

Conflict of Interests
The authors declare no conflict of interests.

Data Availability

No research data are available for this article, as the study is based on publicly accessible online content and qualitative visual interpretation. Meme materials cited are included as figures for illustrative purposes only.

LLMs Disclosure

No LLM tools were used in the preparation of this article.

References

Alonso, L. E. (1994). *La mirada cualitativa en sociología*. Fundamentos.

Arkenbout, C. (2022). Political meme toolkit: Leftist Dutch meme makers share their trade secrets. In C. Arkenbout & L. Scherz (Eds.), *Critical meme reader II* (pp. 20–31). Institute of Network Cultures.

Askanius, T. (2021). On frogs, monkeys, and execution memes: Exploring the humor-hate nexus at the intersection of neo-Nazi and alt-right movements in Sweden. *Television & New Media, 22*(2), 147–165. https://doi.org/10.1177/1527476420982231

Askanius, T., & Keller, N. (2021). Murder fantasies in memes: Fascist aesthetics of death threats and the banalization of white supremacist violence. *Information, Communication & Society, 24*(16), 2522–2539. https://doi.org/10.1080/1369118X.2021.1928273

Bandura, A. (2019). *Principles of behavior modification*. International Psychotherapy Institute.

Barthes, R. (1995). *Lo obvio y lo obtuso: Imágenes, gestos y voces*. Paidós.

Bassil, N., Pourhamzavi, K., & Bayarri, G. (2023). Gramsci, Trotsky and fascism: Tracing the lineage of contemporary right-wing movements in the crisis of capitalism. *Critique, 50*(4), 619–635. https://doi.org/10.1080/03017605.2023.2233100

Bastian, B., & Haslam, N. (2011). Experiencing dehumanization: Cognitive and emotional effects of everyday dehumanization. *Basic and Applied Social Psychology, 33*(4), 295–303. https://doi.org/10.1080/01973533.2011.614132

Bastian, B., Jetten, J., Chen, H., Radke, H. R. M., Harding, J. F., & Fasoli, F. (2013). Losing our humanity: The self-dehumanizing consequences of social ostracism. *Personality and Social Psychology Bulletin, 39*(2), 156–169. https://doi.org/10.1177/0146167212471205

Bayarri, G., Colussi, J., Gomes-Franco e Silva, F., & Arrufat-Martín, S. (2024). Más banderas que palabras: La comunicación no verbal de Bolsonaro en su discurso polarizador. *Anuario Electrónico de Estudios en Comunicación Social "Disertaciones," 17*(1). https://doi.org/10.12804/revistas.urosario.edu.co/disertaciones/a.12802

Bayarri, G., & Fernández-Villanueva, C. (2025). Violence, Hate Speech, and Polarisation in Far-Right Political Influencers in Spain. *Javnost—The Public, 32*(1), 92–114. https://doi.org/10.1080/13183222.2025.2469032

Billig, M. (2019). *Le nationalisme banal*. Presses universitaires de Louvain.

Bowen, B. (2020). "Lol you go to Gulag": The role of sassy socialist memes in Leftbook. *Explorations in Media Ecology, 19*(1), 23–36. https://doi.org/10.1386/eme_00045_1

Brantner, C., Pfurtscheller, D., & Lobinger, K. (2019, May 23–24). *"People only share videos they find entertaining or funny": Right-wing populism, humor and the fictionalization of politics—A case study on the Austrian Freedom Party's 2017 online election campaign videos* [Paper presentation]. Annual Conference of the International Communication Association (ICA), Washington, DC, United States.

Bredekamp, H. (2017). *El acto icónico*. Akal.

Burkart, M., Fraticelli, D., & Palacios, C. (2023). Modalidades de la risa y representaciones de la mujer

inmigrante en *Los amores de Giacumina*. In M. Burkart, D. Fraticelli, & C. Palacios (Eds.), *Arruinando chistes* (pp. 45–60). TeseoPress. https://doi.org/10.55778/TS310005331

Cardoso de Oliveira, L. R. (2004). Honor, dignidad y reciprocidad. *Cuadernos de Antropología Social*, *20*, 25–39.

Cho, J., Ahmed, S., Hilbert, M., Liu, B., & Luu, J. (2020). Do search algorithms endanger democracy? An experimental investigation of algorithm effects on political polarization. *Journal of Broadcasting & Electronic Media*, *64*(2), 150–172. https://doi.org/10.1080/08838151.2020.1757365

Colussi, J., Bayarri, G., & Gomes-Franco e Silva, F. (2023). "We swear to lay down our lives for the fatherland!": Bolsonaro as influencer and agent of political polarization. *Análisis Político*, *106*, 113–132. https://doi.org/10.15446/anpol.v36n106.111044

Denisova, A. (2019). *Internet memes and society: Social, cultural and political contexts*. Routledge.

Dynel, M., & Messerli, T. C. (2020). On a cross-cultural memescape: Switzerland through nation memes from within and from the outside. *Contrastive Pragmatics*, *1*(2), 210–241. https://doi.org/10.1163/26660393-bja10005

Fernández-Villanueva, C., & Bayarri, G. (2021a). Legitimizing hate and political violence through meme images: The Bolsonaro campaign. *Communication and Society*, *4*(2), 449–468. https://doi.org/10.15581/003.34.2.449-468

Fernández-Villanueva, C., & Bayarri, G. (2021b). Nosotros somos héroes, ellos, ni siquiera humanos: Polarización y violencia en la comunicación memética de las extremas derechas española y brasileña. In B. Sánchez-Gutiérrez & A. Pineda (Eds.), *Comunicación política en el mundo digital: Tendencias actuales en propaganda, ideología y sociedad* (pp. 147–180). Dykinson.

Fielitz, M., & Ahmed, R. (2021). *It's not funny anymore: Far-right extremists' use of humour*. Radicalisation Awareness Network. https://home-affairs.ec.europa.eu/system/files/2021-03/ran_ad-hoc_pap_fre_humor_20210215_en.pdf

Fielitz, M., & Thurston, M. (2018). *Post-digital cultures of the far right: Online actions and offline consequences in Europe and the US*. transcript.

Figueruelo Borrieza, A., & Martín Guardado, S. (2023). *Desinformación, odio y polarización*. Aranzadi.

Fraticelli, D. (2023). *El humor hipermediático: Una nueva fase de la mediatización reidera*. Teseo.

Gallardo-Paúls, B. (2018). *Tiempos de hipérbole: Inestabilidad e interferencias en el discurso político*. Tirant Humanidades.

Gomes-Franco e Silva, F., Colussi, J., & Bayarri, G. (2022). El discurso desintermediado de Bolsonaro en Instagram: De ataques a la prensa a simulación de fakenews. In R. Rubira García & A. Labio Bernal (Eds.), *Comunicación, poder y pluralismo cultural. Discursos y desafíos en la esfera pública digital* (pp. 239–264). Fragua.

Gorwa, R. (2019). What is platform governance? *Information, Communication & Society*, *22*(6), 854–871. https://doi.org/10.1080/1369118X.2019.1573914

Hakoköngas, E., Halmesvaara, O., & Sakki, I. (2020). Persuasion through bitter humor: Multimodal discourse analysis of rhetoric in internet memes of two far-right groups in Finland. *Social Media + Society*, *6*(2). https://doi.org/10.1177/2056305120921575

Haslam, N. (2022). Dehumanization and the lack of social connection. *Current Opinion in Psychology*, *43*, 312–316. https://doi.org/10.1016/j.copsyc.2021.08.013

Kien, G. (2019). Memetic irony and the irrationality of political discourse. In L. Shifman (Ed.), *Memes in digital culture* (pp. 123–138). MIT Press.

Mason, L. (2018). *Uncivil agreement: How politics became our identity*. University of Chicago Press.

Merrill, S., Gardell, M., & Lindgren, S. (2024). How 'the Left' meme: Analyzing taboo in the internet memes of r/DankLeft. *New Media & Society*, *26*(1), 45–67. https://doi.org/10.1177/14614448241232144

Merrill, S., & Lindgren, S. (2021). Memes, brands and the politics of post-terror togetherness: Following the Manchester bee after the 2017 Manchester Arena bombing. *Information, Communication & Society*, *24*(16), 2403–2421. https://www.tandfonline.com/doi/full/10.1080/1369118X.2021.1942957

Mey, G., & Dietrich, M. (2016). From text to image: Shaping a visual grounded theory methodology. *Forum Qualitative Social Research*, *17*(2), Article 2. https://doi.org/10.17169/fqs-17.2.2535

Mouffe, C. (2022). *Towards a green democratic revolution: Left populism and the power of affects*. Verso Books.

Parker, I. (2019). *Psychology through critical auto-ethnography*. Routledge.

Parrott, S., & Hopp, T. (2020). Reasons people enjoy sexist humor and accept it as inoffensive. *Atlantic Journal of Communication*, *28*(5), 278–291. https://doi.org/10.1080/15456870.2019.1616737

Pauwels, L. (2015). *Reframing visual social science*. Cambridge University Press.

Pavetich, M., & Stathi, S. (2020). Meta-humanization reduces prejudice, even under high intergroup threat. *Journal of Personality and Social Psychology*, *120*, 651–671. https://doi.org/10.1037/pspi0000259

Pérez, R., & Greene, V. S. (2016). Debating rape jokes vs. rape culture: Framing and counter-framing misogynistic comedy. *Social Semiotics*, *26*(3), 265–282. https://doi.org/10.1080/10350330.2015.1134823

Pinheiro-Machado, R., & Vargas-Maia, T. (Eds.). (2023). *The rise of the radical right in the Global South*. Routledge.

Reynares, J. M. (2024). Las identificaciones políticas en la "nueva derecha" española. Un análisis del caso de VOX desde un enfoque de lógicas. *Política y Sociedad (Madrid)*, *61*(2), Article 87876. https://doi.org/10.5209/poso.87876

Serrano, A., & Zurdo, Á. (2023). *El análisis del discurso en la investigación social: Teorías y prácticas*. Síntesis.

Shifman, L. (2019). Internet memes and the twofold articulation of values. In M. Graham & W. H. Dutton (Eds.), *Society and the internet: How networks of information and communication are changing our lives* (pp. 43–57). Oxford University Press.

Slimovich, A. (2022). *Redes sociales, televisión y elecciones argentinas: La mediatización política en la "era K."* Eudeba.

Soler Gallo, M. (2019). Estrategias persuasivas durante la irrupción de la ideología fascista en España: Miedo, segregación y desprecio por la política. *Círculo de Lingüística Aplicada a la Comunicación*, *80*, 91–114. https://doi.org/10.5209/clac.66602

Wodak, R., Culpeper, J., & Semino, E. (2021). Shameless normalisation of impoliteness: Berlusconi's and Trump's press conferences. *Discourse & Society*, *32*(3), 369–393. https://doi.org/10.1177/0957926520977217

About the Authors

Gabriel Bayarri Toscano is a lecturer at Rey Juan Carlos University. His research focuses on far-right digital culture in Latin America, working at the intersection of anthropology, politics, and communication. He analyzes memetic, affective, and visual strategies, audience responses, and the role of generative AI in political antagonism. He has published widely and held a British Academy Postdoctoral Newton International Fellowship at the University of London.

Concepción Fernández-Villanueva is an Emeritus Professor of social psychology at the Complutense University of Madrid. Her research focuses on the social psychology of violence—particularly gender-based violence and youth violence—examined through media discourse, social representations, and collective identity. She has directed and collaborated on numerous national and international research projects and has published widely in peer-reviewed journals and edited academic volumes.

ARTICLE

The More, the Merrier…: The Effect of Social Network Heterogeneity on Attitudes Toward Political Opponents

Paulina Górska [1] , Dominika Bulska [2,3] , and Maciej Górski [1,4]

[1] Faculty of Psychology, University of Warsaw, Poland
[2] Faculty of Psychology, SWPS University, Poland
[3] Institute for Social Studies, University of Warsaw, Poland
[4] Institute of Psychology, Polish Academy of Sciences, Poland

Correspondence: Dominika Bulska (dominika.bulska@psych.uw.edu.pl)

Abstract

Social network homogeneity is considered one of the key drivers of the rise in affective polarization. As opportunities for contact with different others decrease, out-group animosity increases, fueling political conflict and destabilizing democracy. At the same time, research suggests that diverse social networks foster tolerance for opposing viewpoints. Consistent with the contact hypothesis, empirical studies show that individuals with more politically diverse networks hold more favorable attitudes toward their political opponents. However, it remains unclear whether network heterogeneity affects intergroup relations in the same way as intergroup contact or whether it represents a distinct source of depolarization. Furthermore, there is limited empirical evidence on the psychological mechanisms through which network heterogeneity influences attitudes toward political opponents. In this article, we address these gaps by presenting the results of a 2019 survey study ($N = 378$) conducted in Poland, within a highly polarized context. We show that having a more heterogeneous social network in terms of partisanship is indirectly related to more positive attitudes toward political opponents—an effect explained sequentially by diminished moral conviction and weakened party identification, as well as by weakened party identification alone. Contrary to what has been assumed, this effect is independent of traditionally operationalized intergroup contact, both in terms of its quantity and quality.

Keywords

affective polarization; moralization; party identification; social networks

1. Introduction

Social scientists have long recognized that our social interactions influence who we are. Studies show that people in our social networks often inform our attitudes and behaviors, affecting various aspects of life such as drinking habits (Borsari & Carey, 2001), cultural preferences (e.g., Duricic et al., 2021), or intergroup attitudes (Bracegirdle et al., 2022; Zingora et al., 2020). The influence of friends, family, and associates is also apparent in politics. Empirical evidence suggests that individuals rarely navigate political engagement alone; instead, they rely on close others to gather information, formulate opinions, and make choices about political matters (e.g., Huckfeldt et al., 2004; Lupton & Thornton, 2017; Sinclair, 2012). Importantly, the partisan composition of one's social network may not only shape political preferences and civic engagement but also offer or deny opportunities for meaningful political discussions with those who hold opposing views, which may further translate into attitudes toward political adversaries in general. Consistent with the contact hypothesis (Allport, 1954), empirical research suggests that individuals with more politically diverse social networks have more favorable attitudes toward their political opponents (de Jong, 2024; Ekstrom et al., 2020; Lee, 2022). At the same time, the psychological processes underlying this relationship have yet to be fully explored. Moreover, it remains unclear whether discussion network heterogeneity—defined as the number of close-tie contacts with opposing views and attitudes with whom an individual discusses important matters (Burt, 1984)—influences intergroup relations in the same way as the distant-tie intergroup contact, as traditionally conceptualized in social psychology (see Wölfer & Hewstone, 2017), or whether it represents a qualitatively distinct source of depolarization (see Facciani & Brashears, 2019).

This article addresses these issues by (a) examining moral conviction and party identification as mediators of the relationship between network heterogeneity and affective polarization and (b) comparing the effects of heterogeneous discussion networks with those of self-reported intergroup contact. We propose that diverse discussion networks reduce affective polarization by mitigating the extent to which people perceive political divides as a matter of morality and by limiting party identification. Furthermore, we argue that this sequential mechanism is more likely to explain the depolarizing effect of heterogeneity within one's immediate social environment than the corresponding effects of extra-network intergroup contact, whether in terms of quality or quantity (Islam & Hewstone, 1993).

To verify our claims, we use data from a representative survey of Poles conducted in 2019. In the following sections, we review the extant literature on the link between network heterogeneity and affective polarization, consider moral convictions and party identification as potential mediators, and reflect on the possibly distinct effects of having close vs. distant associates of different political views. After empirically testing our hypotheses, we conclude by situating the present findings within the existing literature and discussing potential avenues for future research.

2. Social Networks and Affective Polarization

Affective political polarization is defined as a tendency to have a positive view of and strong positive feelings toward one's political allies while having a negative view of and strong negative feelings toward one's political opponents (Iyengar & Westwood, 2015). Unlike ideological polarization, which stems from competing views on political issues, affective polarization is primarily based on social identities (Green et al., 2002; Huddy et al., 2015; Huddy & Yair, 2021). In highly polarized societies, affective polarization may take

the form of out-group dehumanization, i.e., denying full humanness to out-group members and people of different political views (Marchlewska et al., 2024; Martherus et al., 2021; Cassese, 2021). In line with social identity theory (Tajfel, 1981), scholars of affective polarization assume that, as a result of the self-categorization process, people divide the world into partisan in-group(s) and political out-group(s). This identification gives rise to in-group bias—a tendency to view the members of one's group more favorably than members of the political out-group—which, in turn, may translate to intergroup hostility (prejudice, dehumanization, or discrimination), especially when one's identity, status, or positive distinctiveness are threatened (Huddy & Yair, 2021). While some degree of policy disagreement in democracies is inevitable or even beneficial, the prevalence of affective polarization is widely considered to be a significant barrier to both functional democracies and effective governance (McCoy & Somer, 2019).

For that reason, social scientists have been trying to identify factors that fuel partisan animosity (for a review, see Iyengar et al., 2019). One of the most prominent explanations suggests that polarization stems from the increasing homogeneity of people's social environments (Enders & Armaly, 2019; Mason, 2016; Zollinger, 2024). For instance, de Jong (2024) demonstrated that adult Americans embedded in more politically heterogeneous discussion networks showed more positive attitudes toward their political opponents. In a similar vein, the analysis of survey data collected in Hong Kong showed that high ego-network difference (i.e., perceived political disagreement between a person and other people located in their discussion network) was associated with lower affective polarization (Lee, 2022).

However, while the negative link between social network heterogeneity and partisan prejudice has already been established, little is known about *why* having a politically diverse discussion network mitigates affective polarization. We propose that weakened moral conviction and reduced party identification—operating both independently and in sequence—may serve as an explanation for this relationship.

3. Attitude Demoralization and In-Group Reappraisal

Growing evidence suggests that politics is often viewed through a moral lens (K. N. Garrett & Bankert, 2020). Intergroup conflict, political contention included, facilitates transforming strong attitudes into moral convictions (see van Zomeren et al., 2024), understood as viewing the attitude object as a reflection of one's fundamental beliefs about what is right and what is wrong (Skitka, 2010). Unlike preferences and conventions, which are rooted in personal taste and in-group norms, respectively, moral convictions are seen as universal and unquestionable truths (Skitka et al., 2021). As a result, individuals with strong moral convictions about an issue often hold negative attitudes toward those expressing divergent opinions (Skitka et al., 2013; Zaal et al., 2017). These effects extend to politics: Those with moral convictions about political issues tend to exhibit greater social distance and prejudice towards out-group partisans (K. N. Garrett & Bankert, 2020; Simonsen & Bonikowski, 2022).

The strength of moral conviction about political issues varies not only between but also within individuals. The process of increasing or decreasing a moral conviction about a given issue is referred to as moralization or demoralization of attitudes, respectively (Skitka et al., 2021). Moralization can occur due to heightened perceptions of harm (Gray et al., 2012; Schein & Gray, 2018), exposure to shocking or emotionally intense attitude-relevant stimuli (Wisneski & Skitka, 2017), collective action participation (Leal et al., 2024), or contact with a disadvantaged out-group (Górska & Tausch, 2023a). Another factor known to facilitate the development

of moralized attitudes is the existence of politically homogeneous networks. In a longitudinal study of Trump and Biden supporters, D'Amore et al. (2024) showed that the increasing perceptions of network homogeneity were followed by the increased moralization of various political topics.

In this contribution, we propose that engaging in meaningful cross-partisan intergroup contact in heterogeneous political discussion networks can lead to de-moralization of politics, ultimately reducing affective polarization. Through such interactions and being exposed to counterattitudinal moral messages (Luttrell et al., 2019), partisans may come to realize that their moral beliefs are not absolute but reflect personal preferences and group conventions—and that others may hold different views (Facciani & Brashears, 2019). Additionally, discussing politics with individuals who hold opposing worldviews may weaken the automatic character of moral reasoning (Baumgartner & Morgan, 2019; Haidt, 2001) and force partisans to actively process their attitudes, thereby de-moralizing them (Ekstrom et al., 2020). Therefore, we expect that greater heterogeneity of discussion networks will be associated with lower moral conviction about politics and, in turn, lower affective polarization (see Figure 1).
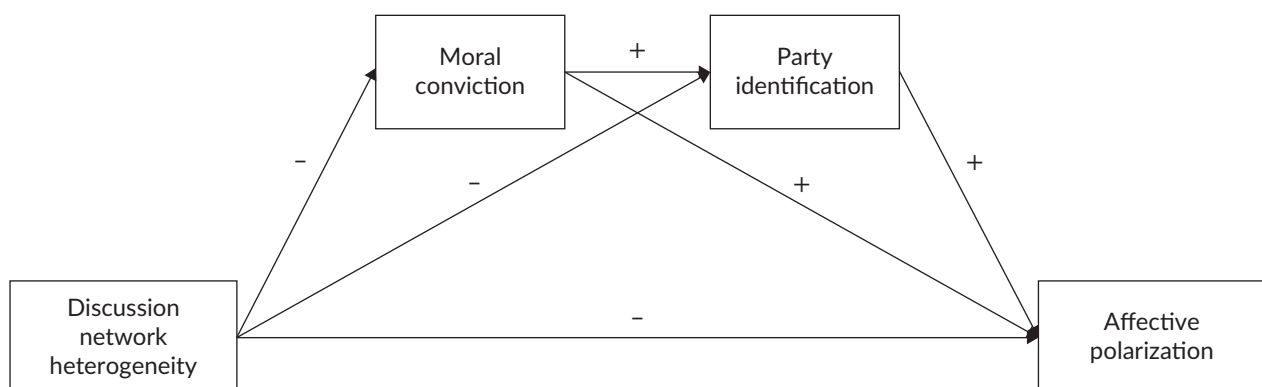
Figure 1. Theoretical model. Note: Controlling for intergroup contact quantity and quality assumed.

Importantly, de-moralization of attitudes due to exposure to cross-cutting beliefs within the discussion network is consistent with the ideas of deprovincialization (Pettigrew et al., 2011; Verkuyten et al., 2022). In contrast to provincialism, defined as "being centered in one's own small world" (Pettigrew, 2011, p. 1) and emerging among segregated and insular groups, deprovincialization results from intergroup contact (Schmid et al., 2013) and encompasses the belief that different cultures may be equally successful in managing the social world. According to its most recent conceptualization (Verkuyten et al., 2022), deprovincialization consists of two interconnected components: in-group reappraisal and out-group openness (also referred to as group and cultural deprovincialization, respectively). The process of in-group reappraisal—originally operationalized as the decrease in in-group identification (e.g., Pettigrew, 2009)—involves recognizing that the in-group's norms and values are not universally shared. Out-group openness, on the other hand, relies on developing greater acceptance and respect for other groups via empathy and perspective-taking.

Although deprovincialization has primarily been studied in the context of interethnic intergroup contact (e.g., Bagci et al., 2024), we believe that heterogeneous discussion networks are likely to have deprovincializing effects in the realm of politics, too, and these effects may be at least partially explained by the de-moralization of politics-related attitudes. The shift from viewing an issue as a matter of morality to seeing it as the function of social norms, which is at the heart of the de-moralization process (Skitka et al.,

2021), seems to overlap with in-group reappraisal, whereby a local character of in-group's norms, values, and traditions is recognized. At the same time, just as politicized identities may develop on the basis and in defense of values cherished by the in-group (Leal et al., 2024; van Zomeren et al., 2024), weakening of moral conviction may result in a decrease in identification with a political in-group. Therefore, it is reasonable to expect that heterogeneous discussion networks would diminish party identification by weakening moral conviction, which, given the identity-based character of affective polarization (Huddy et al., 2015; Huddy & Yair, 2021), may further translate into lower out-party hostility. Hence, we expected a serial mediation from network heterogeneity via moral conviction and party identification to affective polarization (H1).

This is not to suggest that the depolarizing effect of diverse discussion networks operates exclusively through a sequential decrease in moral conviction and party identification (Figure 1). Other pathways and intervening variables are also plausible. For instance, network heterogeneity may strengthen party identification by reducing the proportion or importance of in-group ties, the latter considered an important aspect of identification (e.g., Cameron, 2004). Likewise, the decline in moral conviction may ameliorate affective polarization by dampening negative moral emotions such as contempt, anger, or disgust (Clifford, 2019; Rozin, 1999; van Zomeren et al., 2024), effectively bypassing party identification. Thus, we expect both weakened moral conviction and reduced party identification to independently mediate the negative relationship between discussion network heterogeneity and affective polarization (H2 and H3, respectively).

4. Social Networks and Intergroup Contact

Researchers often attribute the depolarizing effect of heterogeneous discussion networks to intergroup contact (e.g., Facciani & Brashears, 2019; Hobolt et al., 2024). Specifically, being embedded in a politically diverse discussion network implies having intergroup contact with political opponents, while intergroup contact is known to reduce negative out-party attitudes (e.g., Tausch et al., 2024; Wojcieszak & Warner, 2020). While we agree that network heterogeneity and intergroup contact share some characteristics, this overlap is not complete.

We believe that heterogeneous discussion networks—at least as captured by asking individuals whom they discuss important matters with (Burt, 1984)—represent a particular form of intergroup contact, characterized by high levels of self-disclosure, depth, and repeated interaction over time (see Facciani & Brashears, 2019). As shown by Davies et al.'s (2011) meta-analysis on prejudice reduction due to cross-group friendship, these aspects of intergroup relationships yield the strongest effects on out-group-directed attitudes. By contrast, traditionally assessed intergroup contact encompasses a variety of relationships, ranging from superficial interactions with strangers to lifelong friendships (Page-Gould et al., 2022; Pettigrew & Tropp, 2006). Thus, the instances of intergroup contact may differ in quality, depth, and duration.

Another key difference lies in the measurement. Network heterogeneity indices are often based on the division between in-group and out-group (i.e., non-in-group) ties, where the latter serves as the complement of the former. Therefore, the larger proportion of out-group ties translates directly to the smaller proportion of in-group ties. Meanwhile, knowing numerous out-group members—a common operationalization of extensive intergroup contact (e.g., Górska & Tausch, 2023a)—does not necessitate limited relationships within one's in-group.

We believe that these discrepancies may be consequential for the processes of attitude de-moralization and in-group reappraisal. Past research suggests that individuals' definitions of good and evil, as well as their understanding of the world, are shaped through repeated meaningful interactions with close rather than distant associates (e.g., Facciani & Brashears, 2019; Górska & Tausch, 2023a). While immediate social environments formed entirely by like-minded others reinforce belief strength, close ties with dissimilar others weaken one's certainty in the existing beliefs (Berger, 1967; Smith et al., 1999). Building on these findings, one may claim that to serve as an accurate predictor of worldview formation, a measure needs to (a) assess close rather than distant ties and (b) capture the trade-off between in-group and out-group ties. Unlike traditional intergroup contact scales, indices of discussion network heterogeneity meet both these conditions. Therefore, it is reasonable to expect that heterogeneous discussion networks will be a stronger predictor of decreased moral conviction and party identification as compared to intergroup contact in general (H4).

5. Current Research

The present research has two primary goals. First, we aim to examine the psychological mechanisms underlying the impact of social networks on affective polarization, operationalized as in-group bias and blatant dehumanization of political opponents. We propose that the de-moralization of politics and (subsequent) in-group reappraisal will explain the negative relationship between network heterogeneity and affective polarization. Second, we seek to compare the effects of discussion network heterogeneity and more superficial intergroup contact. We hypothesize that repeated, meaningful intergroup contact within discussion networks is more likely to elicit attitude de-moralization and in-group reappraisal processes than intergroup contact beyond one's discussion network.

We test our hypotheses in Poland, a country that has experienced rising popular and elite polarization since the transformation to liberal democracy and free-market capitalism in the early 1990s. In the early post-transformation period, Polish society was primarily polarized with regard to the view on the abolished communist system. In particular, some believed that the former communists should be severely punished and excluded from participation in politics (i.e., lustration), while others proposed that they still should be allowed to sit in the democratic parliament. With most of the former communists retiring or passing, this conflict became less salient. In the mid-2000s, two major parties have emerged: Law and Justice (original name: Prawo i Sprawiedliwość) and Civic Platform (original name: Platforma Obywatelska). A conflict between supporters of these two parties has been growing, especially following the 2010 plane crash in Smolensk, when the death of President Lech Kaczyński (one of the founders of Law and Justice) led to the spread of divisive conspiracy theories of a presumed assassination of the president (Bilewicz et al., 2019). Over time, two antagonistic political forces have solidified—the conservative block, with the leading role of the Law and Justice party (Poland's ruling party from 2015 to 2023), and the pro-democratic block, involving a range of parties opposing the Law and Justice government (which, as the so-called "October 15th coalition," seized power in 2023).

The political conflict between these actors has become increasingly moralized. After winning the elections in 2023, the current Prime Minister Donald Tusk declared: "On October 15th [election day] we defeated pure evil" (Szczęśniak, 2024). Meanwhile, Jarosław Kaczyński, the leader of Law and Justice and brother of President Lech Kaczyński, urged his supporters to "defeat this evil once and for all" during a recent protest

outside the Prime Minister's Office (Bounaoui, 2024). Comparative research (Marchlewska et al., 2024) suggests that the level of affective polarization in Polish society may be close to that found in the United States, despite Poland being a multiparty system, which is generally expected to mitigate polarization (Wagner, 2021). However, most research on affective polarization, including studies on the role of discussion networks, has focused on two-party systems, particularly in the United States (Ekstrom et al., 2020; Facciani & Brashears, 2019). Our study contributes to the literature by examining the mechanisms through which heterogeneous discussion networks and extra-network intergroup contact shape affective polarization in a multiparty context of Poland.

6. Method

We conducted the current research as part of an in-house computer-assisted survey on political and social issues (Górska & Tausch, 2023b; Marchlewska et al., 2022, 2023; Wohl et al., 2020). While the study consisted of three measurements divided by two 6-month intervals, only data from the first wave was used to verify our hypotheses. The reason behind it was the low number of eligible participants (i.e., those who voted for the ruling party or the pro-democratic opposition and provided ego-network data) in the subsequent waves ($n_1 = 378$, $n_2 = 239$, $n_3 = 150$), weakening the statistical power of potential longitudinal analyses.

In the first out of three measurements, a representative sample of adult Poles was achieved using random sampling, with the Polish identity number (PESEL) serving as a sampling frame. The fieldwork was performed by a commercial research company (i.e., Danae) in July and August 2019, a few months ahead of the parliamentary election held on October 13th.

The initial sample consisted of 1,300 individuals (602 men and 698 women), aged between 18 and 93 ($M = 47.20$, $SD = 16.18$). However, as we measured affective polarization only among the ruling party (i.e., Law and Justice) and pro-democratic opposition (i.e., Civic Coalition, Modern, Polish People's Party, the Democratic Left Alliance, Spring, and Together Party) voters, and less than a half (47.8%) of the respondents provided ego-network data, the final sample involved only 378 participants (157 men and 221 women). Respondents' age ranged from 18 to 83 years ($M = 47.80$, $SD = 15.58$). In comparison to the excluded participants, individuals comprising the final sample were better educated, and more likely to be female. No significant differences in terms of age or settlement size were found (see the Supplementary File).

7. Measures

Unless otherwise noted, all measures used the response scale ranging from 1 (*strongly disagree*) to 7 (*strongly agree*). For multi-item measures, the composite scores were obtained by averaging participants' responses.

Participants' *group membership* (i.e., supporting the ruling party or the pro-democratic opposition) was determined based on their response to the following question: "If the parliamentary elections were held this Sunday, for which party candidate would you vote?" Possible responses were the following: 1 = Confederation Liberty and Independence; 2 = Kukiz'15; 3 = Modern; 4 = Together Party: 5 = Civic Coalition; 6 = Polish People's Party; 7 = Law and Justice; 8 = Democratic Left Alliance; 9 = Spring; 10 = Other party; 11 = I don't know; 12 = I will not participate. Participants who indicated support for Law and Justice were classified as ruling party supporters, while those intending to vote for Modern, Together

Party, Civic Coalition, Polish People's Party, Democratic Left Alliance, or Spring were categorized as the *democratic opposition supporters*. Individuals who selected Confederation Liberty and Independence, Kukiz'15, or other party, as well as those who had no voting preference or did not intend to participate in the elections, were not asked about their moral conviction, party identification, or attitudes toward the ruling party and democratic opposition voters. This is because these parties were small and independent from both main polarized blocks. Consequently, they were excluded from the further analyses.

To elicit *ego-network data*, respondents were first asked whether they had talked to anyone in person, on the phone, online, or through any other medium about issues important to them in the past six months. Possible responses were the following: 1 = Yes, I did; 2 = No, I didn't; 3 = I don't know/It's hard to say; 4 = Response refusal. The 6-month timeframe specified in the question matched the 6-month intervals between consecutive measurements (see Hamaker, 2023). Participants who declared talking to anyone were then asked to name up to five individuals (alters) with whom they had discussed important issues during this period (Brashears, 2014). These persons formed participants' discussion networks. For each alter, respondents provided information on various characteristics, including—but not limited to—gender, age, and ties with the remaining alters (for a complete list, see the Supplementary File). The attribute that we used to gauge the heterogeneity of participants' discussion networks were voting preferences, assessed with the following question: "If the parliamentary elections were held this Sunday, for which party candidate do you think [alter's name] would vote?" Possible responses were the following: 1 = Confederation Liberty and Independence; 2 = Kukiz'15; 3 = Modern; 4 = Together Party; 5 = Civic Coalition; 6 = Polish People's Party; 7 = Law and Justice; 8 = Democratic Left Alliance; 9 = Spring; 10 = Other party; 11 = I don't know; 12 = They will not participate. Prior to calculating *network heterogeneity*, we recoded alters' voting preferences into a four-category variable: Alters perceived to support Modern, Together Party, Civic Coalition, Polish People's Party, Democratic Left Alliance, or Spring were categorized as *pro-democratic opposition voters* (1); alters believed to support Law and Justice were categorized as *ruling party* voters (2); alters viewed as the supporters of Confederation Liberty and Independence, Kukiz'15, or other minor parties were categorized as *other party voters* (3); individuals with unspecified political preferences or those unwilling to vote were assigned to the *no preference/non-voters* group (4). We opted for this four-category variable over the original 12-category variable to better reflect the political divisions at the time of data collection (summer 2019) and accurately distinguish between in-group and out-group ties (at the same time, using the original 12-category variable did not change the results in a meaningful way; see the Supplementary File). Moreover, the newly created variable was consistent with the potential development of a superordinate pro-democratic opposition voter identity among the supporters of pro-democratic opposition parties, who may have perceived each other as in-group members rather than out-group members. As all parties categorized here as representing pro-democratic opposition joined the October 15th coalition government after the 2023 parliamentary elections (e.g., Plichta, 2024), treating them as a single block seems justified. Discussion network heterogeneity was operationalized as the E-I index, calculated by subtracting the number of ego's internal ties from external ties and dividing the result by the total number of ties (Krackhardt & Stern, 1988). The index could take values between −1 (the respondent has exclusively in-group ties) to 1 (the respondent has exclusively out-group ties). Participants' E-I scores were obtained with E-NET, a specialized software for the analysis of ego-network data (Borgatti, 2006).

Intergroup contact quantity was measured with two items:

1. "How many supporters of the opposition parties (i.e., Civic Coalition, Modern, Polish People's Party, Democratic Left Alliance, Spring, and Together Party)/the ruling party (i.e., Law and Justice) do you know?" (1 = none; 2 = one; 3 = two; 4 = three; 5 = four; 6 = five or more);
2. "How often do you talk to the supporters of the opposition parties (i.e., Civic Coalition, Modern, Polish People's Party, Democratic Left Alliance, Spring, and Together Party)/the ruling party (i.e., Law and Justice)?" (1 = never; 2 = once a year; 3 = several times per year; 4 = once a month; 5 = once a week; 6 = everyday).

Participants' responses to these questions correlated positively ($r(310) = .51, p < .001$).

To measure *intergroup contact quality*, we asked the respondents how they would describe their relationships with the opposition (i.e., Civic Platform, Modern, Polish People's Party, Democratic Left Alliance, Spring and Left Together?)/the Law and Justice supporters (1 = *negative*, 6 = *positive*; 1 = *unpleasant*, 6 = *pleasant*). Two items creating the quality index correlated positively ($r(346) = .85, p < .001$).

To assess *moral conviction*, we employed two items patterned after the scale provided by van Zomeren et al. (2012): "My opinion about opposition parties' (i.e., Civic Coalition, Modern, Polish People's Party, Democratic Left Alliance, Spring, and Together Party)/ruling party (i.e., Law and Justice) supporters is an important part of my moral norms and values" and "My opinion about opposition parties' (i.e., Civic Coalition, Modern, Polish People's Party, Democratic Left Alliance, Spring, and Together Party)/ruling party (i.e., Law and Justice) reflects an important part of who I am" ($r(360) = .82, p < .001$).

In-group identification was assessed with two items borrowed from Cameron's (2004) scale: "I have a lot in common with other supporters of the opposition parties/ruling party (Law and Justice)" and "Being a supporter of the opposition parties/ruling is an important reflection of who I am" ($r(359) = .63, p < .001$).

To make sure that our conclusions would not depend on the way in which affective polarization is operationalized, we employed two measures of this construct—*blatant dehumanization* and *in-group bias*. *Blatant dehumanization* was assessed with the Ascent of Man pictorial scale (Kteily et al., 2015). Participants viewed an image illustrating five stages of evolutionary progress from ape to human. Alongside the image, a 9-point scale was provided, with 1 representing *the least* and 9 corresponding to *the most* advanced stage of evolution. Participants were asked to indicate how evolved they perceived their in-group and out-group members. The relative dehumanization score was calculated by subtracting out-group humanity rating from in-group humanity rating (e.g., Kteily et al., 2016).

We assessed *in-group bias* with the feeling thermometer (Alwin, 1997), ranging from −50°C (negative feelings) to 50°C (positive feelings). Respondents were asked to indicate their feelings both toward the in-group (either Law and Justice or democratic opposition supporters) as well as the out-group (democratic opposition or Law and Justice supporters, respectively). In-group favoritism score was obtained by subtracting feelings toward the out-group from feelings toward the in-group (e.g., de Jong, 2024). To avoid convergence problems at the analysis stage, the in-group-out-group difference was divided by 10.

To assess participants' *education*, we asked them how many years of full-time education they had completed (an open-ended question). The responses ranged from 3 to 36.

Participants were also asked about the *size of the settlement they lived in*. Possible responses were the following: 1 = rural area; 2 = town up to 50,000 residents; 3 = town between 50,001 and 100,000 residents; 4 = town between 100,001 and 200,000 residents; 5 = city between 200,001 and 500,000 residents; 6 = city with more than 500,000 residents.

8. Results

Table 1 presents descriptive statistics and intercorrelations for the analyzed variables. In comparison to the ruling party supporters, democratic opposition supporters had more heterogeneous discussion networks ($M_{\text{ruling party}} = -0.48$, $SD_{\text{ruling party}} = 0.71$, vs. $M_{\text{democratic opposition}} = -0.27$, $SD_{\text{democratic opposition}} = 0.81$, $t(335.15) = -2.64$, $p = .009$, $d = -0.28$), reported more frequent contact with political opponents ($M_{\text{ruling party}} = 3.64$, $SD_{\text{ruling party}} = 1.41$, vs. $M_{\text{democratic opposition}} = 3.97$, $SD_{\text{democratic opposition}} = 1.50$, $t(339) = -2.10$, $p = .037$, $d = -0.23$), displayed higher level of blatant dehumanization ($M_{\text{ruling party}} = 1.44$, $SD_{\text{ruling party}} = 1.90$, vs. $M_{\text{democratic opposition}} = 2.15$, $SD_{\text{democratic opposition}} = 2.51$, $t(337.40) = -3.05$, $p = .002$, $d = -0.32$), were younger ($M_{\text{ruling party}} = 52.24$, $SD_{\text{ruling party}} = 15.03$, vs. $M_{\text{democratic opposition}} = 43.37$, $SD_{\text{democratic opposition}} = 14.87$, $t(376) = 5.77$, $p < .001$, $d = 0.59$) and better-educated ($M_{\text{ruling party}} = 12.64$, $SD_{\text{ruling party}} = 3.55$, vs. $M_{\text{democratic opposition}} = 13.92$, $SD_{\text{democratic opposition}} = 2.97$, $t(366) = -3.75$, $p < .001$, $d = 0.39$), and lived in larger settlements ($M_{\text{ruling party}} = 2.22$, $SD_{\text{ruling party}} = 1.66$, vs. $M_{\text{democratic opposition}} = 3.07$, $SD_{\text{democratic opposition}} = 1.86$, $t(371.18) = -4.66$, $p < .001$, $d = -0.48$). The two groups did not differ in terms of intergroup contact quality, moral conviction, party identification, in-group bias and gender distribution (all ps > .101).

Table 1. Descriptive statistics and intercorrelations for ruling party and opposition supporters.

		M	SD	1	2	3	4	5	6	7	8	9	10
1.	Ego-network heterogeneity	−0.38	0.76	—									
2.	Intergroup contact quality	3.77	1.29	.16**	—								
3.	Intergroup contact quantity	3.81	1.46	.20***	.33***	—							
4.	Moral conviction	4.69	1.62	−.13*	−.04	.14**	—						
5.	Party identification	4.85	1.47	−.16**	.03	.29***	.55***	—					
6.	Blatant dehumanization	1.80	2.25	−.04	−.34***	−.02	.13*	.19***	—				
7.	In-group bias	4.34	3.09	−.20***	−.45***	−.04	.21***	.29***	.65***	—			
8.	Gender (0 = F, 1 = M)	0.42	0.49	.10	−.06	−.03	−.003	−.04	−.04	−.05	—		
9.	Age	47.80	15.58	−.07	−.11*	−.08	.09	.04	.06	.15**	.04	—	
10.	Years of education	13.29	3.33	.05	.03	.08	−.07	−.04	−.02	.02	−.13*	−.34***	—
11.	Settlement size	2.65	1.81	.05	−.07	−.02	−.09	−.08	−.04	.02	.03	−.07	.21***

Notes: *** $p < .001$. ** $p < .01$. * $p < .05$; correlation coefficients for the ruling party and democratic opposition supporters are presented above and below the diagonal, respectively.

9. Hypotheses Testing

We verified our theorizing in the structural equation modelling (SEM) framework. To simultaneously examine hypotheses 1–4 among the ruling party and democratic opposition voters, we conducted a multiple-group analysis using observed variables. While *discussion network heterogeneity, intergroup contact quantity,* and *intergroup contact quality* were specified as exogenous predictors, moral conviction and party identification were defined as consecutive mediators, and blatant dehumanization and in-group bias served as the dependent variables (see Figure 2). The analyses used the maximum likelihood estimation and were evaluated according to the established fit indices criteria (Hu & Bentler, 1998). Statistical significance of indirect effects was determined based on 95% CIs obtained with bootstrapping (5,000 samples). The full information maximum likelihood method was employed to handle missing data. All analyses were performed in lavaan (Rosseel, 2012) using R.
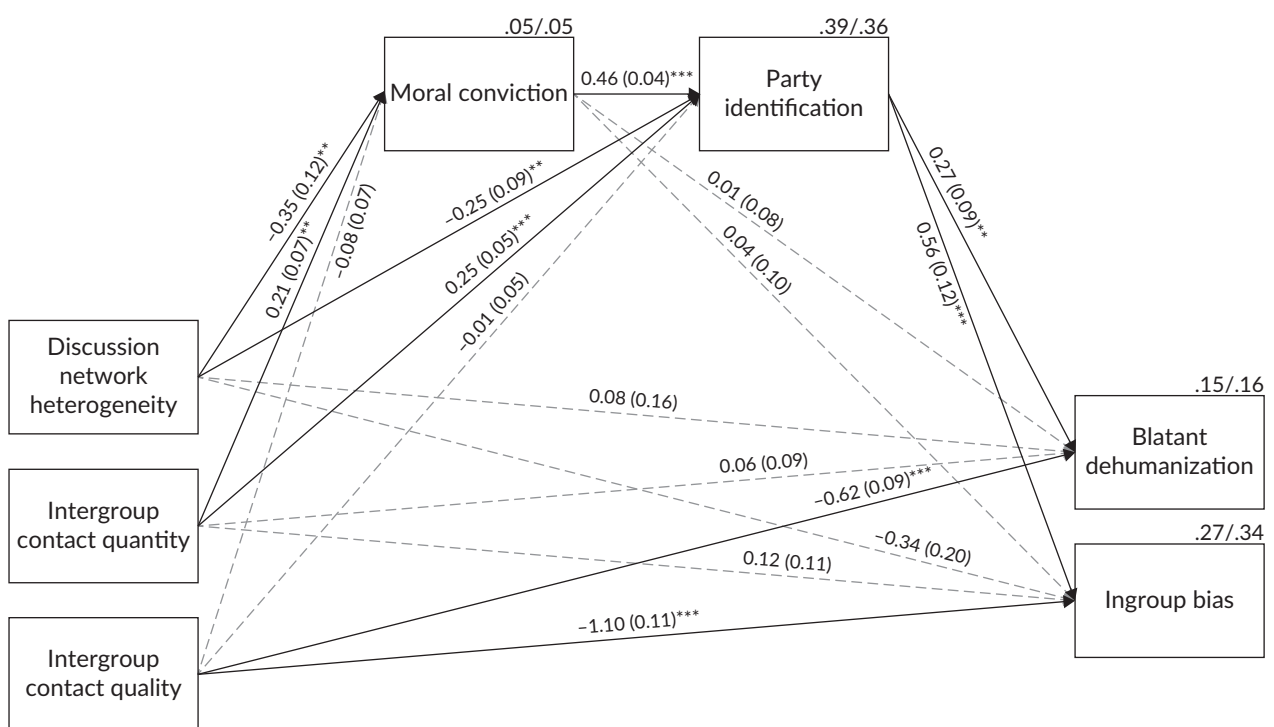
Figure 2. Structural and psychological antecedents of affective polarization. Notes: Entries are unstandardized estimates; solid and dashed lines represent significant and nonsignificant effects, respectively; R^2 coefficients for the ruling party supporters are presented first.

Model 1, which imposed no equality constraints across the two groups, was saturated and, by definition, fitted the data perfectly. At the same time, it did not outperform Model 2 that forced the corresponding paths to be equal for the ruling party and pro-democratic opposition supporters, $\chi^2(21) = 24.79$, $p = .256$, CFI = .993, RMSEA = .031, SRMR = .064, $\Delta\chi^2(21) = 24.79$, $p = .256$. Therefore, we assumed that the investigated processes were the same across the two groups and interpreted the results of a more parsimonious Model 2 (Figure 2).

In line with our expectations, discussion network heterogeneity was a negative predictor of moral conviction, moral conviction served as a positive predictor of party identification, and party identification

predicted blatant dehumanization and in-group bias positively. The serial indirect effect of ego-network heterogeneity on blatant dehumanization via reduced moral conviction and party identification was negative and significant ($IE = -0.04$, $SE = 0.02$, 95% CI [−0.10, −0.01]). In a similar vein, discussion network heterogeneity was indirectly related to lower in-group bias; this relationship was sequentially mediated by decreased moral conviction and party identification ($IE = -0.89$, $SE = 0.39$, 95% CI [−1.74, −0.23]). Thus, irrespective of how the affective polarization was operationalized, the results consistently supported H1.

Notably, the pathway through moral conviction and party identification was not the only mechanism linking discussion network heterogeneity and the outcome variables. Decreased party identification independently mediated the relationships between discussion network heterogeneity and both blatant dehumanization ($IE = -0.07$, $SE = 0.04$, 95% CI [−0.15, −0.01]) and in-group bias ($IE = -1.39$, $SE = 0.57$, 95% CI [−2.65, −0.39]). By contrast, reduced moral conviction alone did not constitute a significant pathway to blatant dehumanization ($IE = -0.005$, $SE = 0.03$, 95% CI [−0.07, 0.06]) or in-group bias ($IE = -0.10$, $SE = 0.36$, 95% CI [−0.87, 0.60]). These results were consistent with H3 but not with H2.

To verify H4, according to which the processes of de-moralization and in-group reappraisal are more likely to be triggered by discussion network heterogeneity than intergroup contact beyond one's immediate social environment, we compared the exogenous variables' effects on moral conviction and party identification. In line with our expectations, the negative effect of discussion network heterogeneity on moral conviction differed significantly from the corresponding *positive* effect of intergroup contact quantity ($\Delta\chi^2(1) = 16.25$, $p < .001$), while the difference with the null effect of intergroup contact quality was on the verge of significance ($\Delta\chi^2(1) = 3.69$, $p = .055$). In a similar vein, the total effect of discussion network heterogeneity on party identification was negative ($TE = -0.42$, $SE = 0.10$, 95% CI [−0.59, −0.23]), and differed significantly from the positive total effect of intergroup contact quantity ($TE = 0.35$, $SE = 0.06$, 95% CI [0.23, 0.46]; $\Delta TE = -0.76$, $SE = 0.12$, 95% CI [−0.99, −0.55]) and the null total effect of intergroup contact quality ($TE = -0.04$, $SE = 0.06$, 95% CI [−0.19, 0.10]; $\Delta TE = -0.38$, $SE = 0.13$, 95% CI [−0.61, −0.14]). Thus, present data was rather supportive of H4.

Importantly, the unexpected, positive effects of intergroup contact quantity on moral conviction and party identification translated to the level of affective polarization. In particular, the quantity of contact with political opponents indirectly predicted stronger blatant dehumanization and in-group bias by the sequence of increased moral conviction and party identification ($IE = 0.03$, $SE = 0.01$, 95% CI [0.01, 0.05] and $IE = 0.54$, $SE = 0.20$, 95% CI [0.20, 0.97] respectively). Additionally, the relationship between intergroup contact quantity and both outcome variables was mediated by increased party identification alone ($IE = 0.07$, $SE = 0.03$, 95% CI [0.02, 0.12] for blatant dehumanization and $IE = 1.38$, $SE = 0.40$, 95% CI [0.68, 2.24] for in-group bias). On the other hand, neither moral conviction nor party identification explained the negative effects of intergroup contact quality on blatant dehumanization and in-group bias (see Figure 2).

10. Robustness Checks

To assess the robustness of current findings, we repeated the analyses using alternative operationalizations of discussion network composition. Specifically, we used the original 12-category variable to represent party preferences as well as the proportion of alters showing the same party preference as ego, which has been used as the measure of discussion network homogeneity (the opposite of heterogeneity) in past research

(e.g., Facciani & Brashears, 2019). These modifications led to only minor changes in our conclusions (see the Supplementary File).

11. Additional Analyses

Although our theorizing posits a specific structure of connections between discussion network heterogeneity, moral conviction, party identification, and affective polarization, it should be acknowledged that other patterns of relationships between these variables are also theoretically plausible. For instance, it is legitimate to expect that strong moral conviction would moderate the effects of discussion network heterogeneity on party identification and out-group dislike (Figure 3).
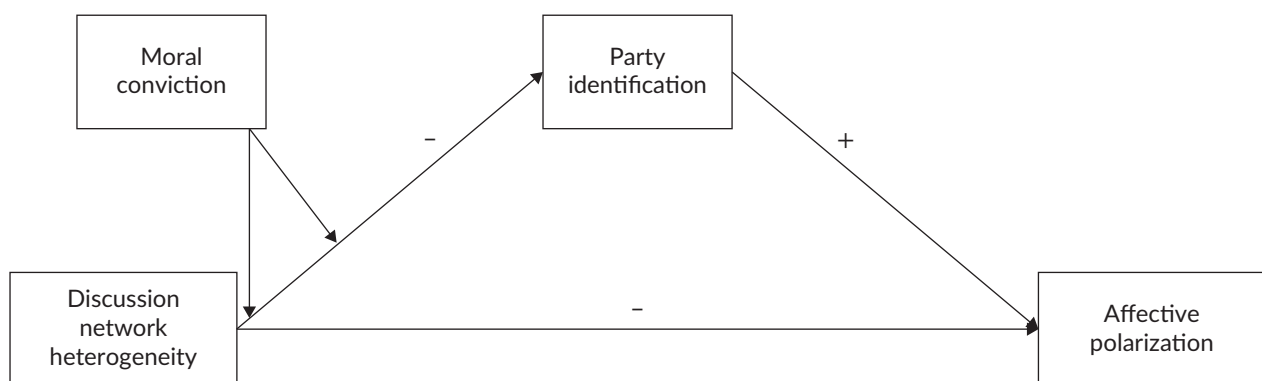
Figure 3. Alternative theoretical model. Note: Controlling for intergroup contact quantity and quality is assumed.

As revealed in the past research, due to their absolute character, high moral convictions diminish associations between typically related phenomena (de Cristofaro et al., 2021). Therefore, the negative effects of discussion network heterogeneity on affective polarization (both the indirect effect via decreased party identification and the direct effect) may be weak or non-existent at high levels of moral conviction and emerge only when moral conviction is low. Alternatively, if network heterogeneity works in a similar manner as intergroup contact, its prejudice-reduction effects may be the strongest among those partisans who are most strongly predisposed to show high initial levels of out-group hostility (e.g., Dhont & van Hiel, 2009, 2011; Hodson et al., 2009). Given that moral conviction correlates positively with affective polarization (K. N. Garrett & Bankert, 2020), the negative effect of discussion network heterogeneity should be most strongly pronounced at high levels of moral conviction. Additional analyses indicated that our data were more supportive of this latter possibility (see the Supplementary File).

12. Discussion

This study had two objectives. First, we sought to uncover the psychological mechanisms that explain the relationship between discussion network heterogeneity and affective polarization. Second, we investigated whether heterogeneous discussion networks and extra-network intergroup contact (over and above the contact one has in their network) have distinct effects on attitudes toward political opponents and their psychological antecedents, specifically moral conviction and party identification.

The analysis of survey data from the Polish political context supported our hypotheses. A greater presence of out-group ties in one's discussion network was indirectly associated with reduced in-group bias and lower blatant dehumanization of political opponents. These effects were mediated by a sequence of diminished moral conviction and weakened party identification, as well as by lower party identification alone. In contrast, the quantity of intergroup contact beyond one's ego network was positively related to in-group bias and out-group dehumanization. These effects were explained by heightened moral conviction and party identification operating in sequence, as well as by the increased party identification on its own. At the same time, high-quality intergroup contact beyond one's discussion network was linked to lower in-group bias and blatant dehumanization, though these effects were not explained by the changes in moral conviction and party identification.

Our findings contribute to the literature in several ways. First, we provide evidence on the mechanisms underpinning the negative link between discussion network heterogeneity and affective polarization. While past research has shown that being embedded in a politically diverse social environment is associated with lower dislike of political opponents (e.g., de Jong, 2024; Lee, 2022), it had not been clear *why* this was the case. Our study addressed this gap by identifying two key psychological processes—one involving the sequence of attitude de-moralization and in-group reappraisal, and the other relying on the independent decline in party identification—as responsible for the depolarizing effect of a heterogeneous discussion network. As such, current results support the proposition that the link between social structure and political conflict (de)escalation is at least partially accounted for by the moralization and identification processes (d'Amore et al., 2024; van Zomeren et al., 2024).

Next, present findings highlight key differences in the effects of close and distant outparty ties. While previous studies have examined the impact of close ties to out-partisans (understood as cross-cutting integroup contact in one's discussion network) on ideological (e.g., Facciani & Brashears, 2019) and affective polarization (R. K. Garrett et al., 2014; Lee, 2022; Robinson, 2010), research directly contrasting the effects of close and distant ties remains scarce. To our knowledge, only one study has made such a comparison, showing that both heterogeneous discussion networks (i.e., close ties) and acquaintance networks (i.e., distant ties) predict lower partisan animosity, with the latter having a slightly stronger effect (de Jong, 2024). Our findings expand on this evidence by taking a more nuanced perspective on intergroup contact beyond the discussion networks—differentiating between contact quantity and quality—and investigating how close and distant outparty ties relate to psychological processes underlying political polarization. Notably, we show that even at the correlational level (see Table 1), two quantity-focused operationalizations of interparty relationships—discussion network heterogeneity and intergroup contact quantity—exhibit opposite relationships with moral conviction and party identification. Specifically, while discussion network heterogeneity is linked to attitude de-moralization and in-group reappraisal, greater contact quantity is related to stronger moralization and in-group identification. These results are in line with the suggestion that close out-group members are more important for weakening one's political beliefs than the distant ones (Facciani & Brashears, 2019).

At the same time, although high-quality intergroup contact beyond the discussion network shows strong, negative relationships with different measures of affective polarization, these effects are not explained by changes in moral conviction or in-group identification This suggests that while positive intergroup contact with the out-partisans beyond the closest social milieu fosters more favorable attitudes toward political

opponents—perhaps even more effectively than discussion network heterogeneity (de Jong, 2024)—it does not necessarily trigger attitude de-moralization in in-group reappraisal processes. Consistent with past theorizing (Berger, 1967; Smith et al., 1999) and empirical findings (Facciani & Brashears, 2019), perceiving the political divisions through a less moralized lens and developing a more nuanced view of the in-group seem to require interactions with close outparty associates rather than pleasant encounters with distant out-partisans.

Finally, current results add to the literature by showing that some aspects of non-negative inter-party contact may be related to *greater* affective polarization. Specifically, when discussion network heterogeneity and intergroup contact quality are controlled for, the quantity of extra-network intergroup contact serves as a positive predictor of moral conviction and in-group identification, which, in turn, are associated with greater out-group dislike. This result is consistent with past research showing that social identity threat reverses the positive relationship between the repeated exposure to out-group-relevant stimuli and positive emotions towards those objects (Crisp et al., 2009). In a similar vein, given that contact with political opponents is often perceived as threatening (see Wojcieszak & Warner, 2020), frequent interactions with distant out-party associates may enhance partisan animosity.

Despite these contributions, our research has several limitations, most of which stem from suboptimal measurement. First, due to the high cost of face-to-face data collection, we used a numerically constrained name generator, which may have introduced an error to the measures of network composition (Bidart & Charbonneau, 2011; see also Kogovšek & Hlebec, 2005). The five-alter upper limit may have failed to capture some influential alters, the presence of whom would change participants' discussion network heterogeneity scores. On the other hand, 84.7% participants named from 1 to 4 alters, which suggests that insufficient coverage was a rather minor problem in the present sample.

Second, network homogeneity was assessed based on participants' self-reports collected through face-to-face interviews. This mode of data collection comes with a range of potential limitations, such as recall bias (Bell et al., 2007), social desirability bias (Nederhof, 1985), and inaccuracies in participants' perception of alters' attributes or alter-alter ties (e.g., Goel et al., 2010). Replicating present findings using whole-network data would substantially strengthen confidence in the robustness of these results.

Next, our measures of intergroup contact were rather general and did not explicitly distinguish between close and distant out-group associates. Given their weak correlations with discussion network heterogeneity (see Table 1), we assumed that intergroup contact quantity and quality scales captured distant rather than close ties. Additionally, by incorporating both intergroup contact and discussion heterogeneity into the same structural model, we controlled for the variance that broad intergroup contact shared with close out-group ties comprising one's discussion network. This approach allowed us to interpret the effects of contact quantity and quality as reflecting interactions with more distant out-group associates. However, future studies should directly assess contact with political opponents beyond one's close ties—for instance, by explicitly inquiring about interactions with out-group representatives different than those identified using the important matters name generator (see de Jong, 2024).

Then, while current research suggests that contact-driven deprovincialization extends to political context, it overlooks the multifaceted character of this construct. Early accounts of the deprovincialization hypothesis

(Pettigrew, 1997, 2009) framed it primarily as an in-group-related phenomenon. However, more recent conceptualizations suggest that, apart from in-group reappraisal, deprovincialization includes an out-group-related dimension—namely, increased out-group openness (Verkuyten et al., 2022). By operationalizing deprovincialization as decreased party identification, current research neglects this out-group-related component. This omission may be important, as, compared to in-group reappraisal, out-group openness is more strongly linked to intergroup experiences, such as intergroup contact (e.g., Boin et al., 2020; Lucarini et al., 2023). It is therefore conceivable that high-quality intergroup contact—either within or beyond one's discussion network—mitigates affective polarization not only by shifting one's perspective on the in-group, as the present results suggest, but also by expanding one's openness to alternative worldviews. This mechanism may be particularly relevant for intergroup contact quality, which predicted lower affective polarization directly, but not through moral conviction or party identification. To explore this possibility, future research would do well to adopt a more granular view of deprovincialization.

At the same time, other psychological processes may also explain why high-quality contact with distant outparty associates is related to lower affective polarization. The list of candidate mechanisms involves weaker intergroup anxiety and stronger out-group-directed empathy (Pettigrew & Tropp, 2008), higher commonality perceptions (Wojcieszak & Warner, 2020), or more favorable metaperceptions (Landry et al., 2023; Moore-Berg et al., 2020), to name just a few. Assessing all these mechanisms simultaneously would grant the opportunity to determine their relative importance and, perhaps, identify further differences in the psychological consequences of close and distant out-group ties.

Finally, like any cross-sectional study, the current research does not warrant strong causal inference. While the present results align with our expectations regarding the relative ordering of variables in the causal chain, they do not rule out alternative patterns of relationships (e.g., discussion network composition being a consequence rather than a cause of affective polarization). Thus, we advise caution in interpreting these findings and recommend their replication in an adequately powered longitudinal research.

Acknowledging all these limitations, we believe the present findings offer unique insights into the depolarizing potential of heterogeneous discussion networks and encourage further integration between social networks, intergroup contact, and affective polarization literatures.

Conflict of Interests
The authors declare no conflict of interests.

Data Availability
Data and code necessary to replicate our analyses are available here: https://osf.io/zakws/?view_only=0694ab73eb3a4e1e9df5c716ae13344d.

LLMs Disclosure
The first draft of this article was refined for clarity and language with the assistance of ChatGPT.

Supplementary Material

Supplementary material for this article is available online in the format provided by the author (unedited).

References

Allport, G. W. (1954). *The nature of prejudice*. Addison Wesley.

Alwin, D. F. (1997). Feeling thermometers versus 7-point scales: Which are better? *Sociological Methods & Research*, *25*(3), 318–340. https://doi.org/10.1177/0049124197025003003

Bagci, S. C., Çoksan, S., Turnuklu, A., & Tercan, M. (2024). Do positive and negative intergroup contact create shifts in in-group and out-group attitudes over time: A three-wave longitudinal study testing alternative mediation models. *Journal of Community & Applied Social Psychology*, *34*(6), Article 70011. https://doi.org/10.1002/casp.70011

Baumgartner, J. N., & Morgan, G. S. (2019). Mindfulness and cognitive depletion shape the relationship between moral conviction and intolerance of dissimilar others. *Studia Psychologica*, *61*(1), 31–41. https://doi.org/10.21909/sp.2019.01.770

Bell, D. C., Belli-McQueen, B., & Haider, A. (2007). Partner naming and forgetting: Recall of network members. *Social Networks*, *29*(2), 279–299. https://doi.org/10.1016/j.socnet.2006.12.004

Berger, P. L. (1967). *The sacred canopy: Elements of a sociological theory of religion*. Anchor Books.

Bidart, C., & Charbonneau, J. (2011). How to generate personal networks: Issues and tools for a sociological perspective. *Field Methods*, *23*(3), 266–286. https://doi.org/10.1177/1525822X11408513

Bilewicz, M., Witkowska, M., Pantazi, M., Gkinopoulos, T., & Klein, O. (2019). Traumatic rift: How conspiracy beliefs undermine cohesion after societal trauma? *Europe's Journal of Psychology*, *15*(1), 82–93. https://doi.org/10.5964/ejop.v15i1.1699

Boin, J., Fuochi, G., & Voci, A. (2020). Deprovincialization as a key correlate of ideology, prejudice, and intergroup contact. *Personality and Individual Differences*, *157*, Article 109799. https://doi.org/10.1016/j.paid.2019.109799

Borgatti, S. P. (2006). *E-NET software for the analysis of ego−Network data* [Software]. Analytic Technologies. https://sites.google.com/site/enetsoftware1

Borsari, B., & Carey, K. B. (2001). Peer influences on college drinking: A review of the research. *Journal of Substance Abuse*, *13*(4), 391–424.

Bounaoui, S. (2024, January 11). Kaczyński przed kancelarią Tuska. "Raz na zawsze pokonać to zło." *Wirtualna Polska*. https://wiadomosci.wp.pl/walka-nie-bedzie-latwa-kaczynski-apelowal-przed-kancelaria-premiera-6983662039538208a

Bracegirdle, C., Reimer, N. K., van Zalk, M., Hewstone, M., & Wölfer, R. (2022). Disentangling contact and socialization effects on out-group attitudes in diverse friendship networks. *Journal of Personality and Social Psychology*, *122*(1), 1–15. https://doi.org/10.1037/pspa0000240

Brashears, M. E. (2014). "Trivial" topics and rich ties: The relationship between discussion topic, alter role, and resource availability using the "important matters" name generator. *Sociological Science*, *1*, 493–511. https://doi.org/10.15195/v1.a27

Burt, R. S. (1984). Network items and the General Social Survey. *Social Networks*, *6*(4), 293–339. https://doi.org/10.1016/0378-8733(84)90007-8

Cameron, J. E. (2004). A three-factor model of social identity. *Self and Identity*, *3*(3), 239–262. https://doi.org/10.1080/13576500444000047

Cassese, E. C. (2021). Partisan dehumanization in American politics. *Political Behavior*, *43*(1), 29–50. https://doi.org/10.1007/s11109-019-09545-w

Clifford, S. (2019). How emotional frames moralize and polarize political attitudes. *Political Psychology*, *40*(1), 75–91. https://doi.org/10.1111/pops.12507

Crisp, R. J., Hutter, R. R., & Young, B. (2009). When mere exposure leads to less liking: The incremental threat effect in intergroup contexts. *British Journal of Psychology*, *100*(1), 133–149. https://doi.org/10.1348/000712608X318635

D'Amore, C., van Zomeren, M., & Koudenburg, N. (2024). How perceived polarization predicts attitude moralization (and vice versa): A four-wave longitudinal study during the 2020 U.S. election. *Journal of Personality and Social Psychology*, *126*(4), 624–642. https://doi.org/10.1037/pspi0000454

Davies, K., Tropp, L. R., Aron, A., Pettigrew, T. F., & Wright, S. C. (2011). Cross-group friendships and intergroup attitudes: A meta-analytic review. *Personality and Social Psychology Review*, *15*(4), 332–351. https://doi.org/10.1177/1088868311411103

de Cristofaro, V., Pellegrini, V., Giacomantonio, M., Livi, S., & Zomeren, M. (2021). Can moral convictions against gender inequality overpower system justification effects? Examining the interaction between moral conviction and system justification. *British Journal of Social Psychology*, *60*(4), 1279–1302. https://doi.org/10.1111/bjso.12451

de Jong, J. F. (2024). *Enduring divides? Social networks and the entrenchment of political polarization* [Unpublished doctoral dissertation]. European University Institute.

Dhont, K., & van Hiel, A. (2009). We must not be enemies: Interracial contact and the reduction of prejudice among authoritarians. *Personality and Individual Differences*, *46*(2), 172–177. https://doi.org/10.1016/j.paid.2008.09.022

Dhont, K., & van Hiel, A. (2011). Direct contact and authoritarianism as moderators between extended contact and reduced prejudice: Lower threat and greater trust as mediators. *Group Processes & Intergroup Relations*, *14*(2), 223–237. https://doi.org/10.1177/1368430210391121

Duricic, T., Kowald, D., Schedl, M., & Lex, E. (2021). My friends also prefer diverse music: Homophily and link prediction with user preferences for mainstream, novelty, and diversity in music. In M. Coscia, A. Cuzzocrea, & K. Shu (Eds.), *Proceedings of the 2021 IEEE/ACM international conference on advances in social networks analysis and mining* (pp. 447–454). https://doi.org/10.1145/3487351.3492706

Ekstrom, P. D., Smith, B. A., Williams, A. L., & Kim, H. (2020). Social network disagreement and reasoned candidate preferences. *American Politics Research*, *48*(1), 132–154. https://doi.org/10.1177/1532673X19858343

Enders, A. M., & Armaly, M. T. (2019). The differential effects of actual and perceived polarization. *Political Behavior*, *41*, 815–839. https://doi.org/10.1007/s11109-018-9476-2

Facciani, M., & Brashears, M. E. (2019). Sacred alters: The effects of ego network structure on religious and political beliefs. *Socius*. Advance online publication. https://doi.org/10.1177/2378023119873825

Garrett, K. N., & Bankert, A. (2020). The moral roots of partisan division: How moral conviction heightens affective polarization. *British Journal of Political Science*, *50*(2), 621–640. https://doi.org/10.1017/S000712341700059X

Garrett, R. K., Dvir Gvirsman, S., Johnson, B. K., Tsfati, Y., Neo, R., & Dal, A. (2014). Implications of pro- and counterattitudinal information exposure for affective polarization. *Human Communication Research*, *40*(3), 309–332. https://doi.org/10.1111/hcre.12028

Goel, S., Mason, W., & Watts, D. J. (2010). Real and perceived attitude agreement in social networks. *Journal of Personality and Social Psychology*, *99*(4), 611–621. https://doi.org/10.1037/a0020697

Górska, P., & Tausch, N. (2023a). Dynamic, yet stable: Separating within-and between-person components of collective action in support of a disadvantaged out-group and its antecedents. *Social Psychological and Personality Science*, *14*(7), 875–887. https://doi.org/10.1177/19485506221133882

Górska, P., & Tausch, N. (2023b). People or ideology? Social conservatism and intergroup contact moderate heterosexuals' responses to a state-sponsored anti-LGBT campaign. *Sexuality Research and Social Policy*, *20*(3), 1049–1063. https://doi.org/10.1007/s13178-022-00783-y

Gray, K., Young, L., & Waytz, A. (2012). Mind perception is the essence of morality. *Psychological Inquiry*, *23*(2), 101–124. https://doi.org/10.1080/1047840X.2012.651387

Green, D., Palmquist, B., & Schickler, E. (2002). *Partisan hearts and minds: Political parties and the social identities of voters*. Yale University Press.

Haidt, J. (2001). The emotional dog and its rational tail: a social intuitionist approach to moral judgment. *Psychological Review*, *108*(4), Article 814. https://doi.org/10.1037/0033-295x.108.4.814

Hamaker, E. L. (2023). The within-between dispute in cross-lagged panel research and how to move forward. *Psychological Methods.* Advance online publication. https://doi.org/10.1037/met0000600

Hobolt, S. B., Lawall, K., & Tilley, J. (2024). The polarizing effect of partisan echo chambers. *American Political Science Review*, *118*(3), 1464–1479. https://doi.org/10.1111/j.1467-9221.2004.00357.x

Hodson, G., Harry, H., & Mitchell, A. (2009). Independent benefits of contact and friendship on attitudes toward homosexuals among authoritarians and highly identified heterosexuals. *European Journal of Social Psychology*, *39*(4), 509–525. https://doi.org/10.1002/ejsp.558

Hu, L.-t., & Bentler, P. M. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychological Methods*, *3*(4), 424–453. https://doi.org/10.1037/1082-989X.3.4.424

Huckfeldt, R., Mendez, J. M., & Osborn, T. (2004). Disagreement, ambivalence, and engagement: The political consequences of heterogeneous networks. *Political Psychology*, *25*(1), 65–95. https://doi.org/10.1111/j.1467-9221.2004.00357.x

Huddy, L., Mason, L., & Aarøe, L. (2015). Expressive partisanship: Campaign involvement, political emotion, and partisan identity. *American Political Science Review, 109*(1), 1–17. https://doi.org/10.1017/S0003055414000604

Huddy, L., & Yair, O. (2021). Reducing affective polarization: Warm group relations or policy compromise? *Political Psychology*, *42*(2), 291–309. https://doi.org/10.1111/pops.12699

Islam, M. R., & Hewstone, M. (1993). Intergroup attributions and affective consequences in majority and minority groups. *Journal of Personality and Social Psychology*, *64*(6), 936–950. https://doi.org/10.1037/0022-3514.64.6.936

Iyengar, S., Lelkes, Y., Levendusky, M., Malhotra, N., & Westwood, S. J. (2019). The origins and consequences of affective polarization in the United States. *Annual Review of Political Science*, *22*(1), 129–146. https://doi.org/10.1146/annurev-polisci-051117-073034

Iyengar, S., & Westwood, S. J. (2015). Fear and loathing across party lines: New evidence on group polarization. *American Journal of Political Science*, *59*(3), 690–707. https://doi.org/10.1111/ajps.12152

Kogovšek, T., & Hlebec, V. (2005). Effects of limitation of number of alters and time frame in the Burt name generator. *Advances in Methodology and Statistics*, *2*(1), 59–71.

Krackhardt, D., & Stern, R. N. (1988). Informal networks and organizational crises: An experimental simulation. *Social Psychology Quarterly*, *51*(2), 123–140. https://doi.org/10.2307/2786835

Kteily, N., Bruneau, E., Waytz, A., & Cotterill, S. (2015). The ascent of man: Theoretical and empirical evidence for blatant dehumanization. *Journal of Personality and Social Psychology*, *109*(5), 901–931. https://doi.org/10.1037/pspp0000048

Kteily, N., Hodson, G., & Bruneau, E. (2016). They see us as less than human: Metadehumanization predicts intergroup conflict via reciprocal dehumanization. *Journal of Personality and Social Psychology*, *110*(3), 343–370. https://doi.org/10.1037/pspa0000044

Landry, A. P., Schooler, J. W., Willer, R., & Seli, P. (2023). Reducing explicit blatant dehumanization by correcting exaggerated meta-perceptions. *Social Psychological and Personality Science*, *14*(4), 407–418. https://doi.org/10.1177/19485506221099146

Leal, A., van Zomeren, M., González, R., Gordijn, E., Carozzi, P., Reifen-Tagar, M., Álvarez, B., Frigolett, C., & Halperin, E. (2024). Attitude moralization in the context of collective action: How participation in collective action may foster moralization over time. *Journal of Personality and Social Psychology*. Advance online publication. https://doi.org/10.1037/pspi0000486

Lee, F. L. (2022). Ego-network difference, political communication, and affective polarization during political contention. *International Journal of Communication*, *16*, Article 24.

Lucarini, A., Boin, J., Fuochi, G., Voci, A., Verkuyten, M., & Pettigrew, T. F. (2023). The nature of deprovincialism: Assessment, nomological network, and comparison of cultural and group deprovincialization. *Journal of Community & Applied Social Psychology*, *33*(4), 868–881. https://doi.org/10.1002/casp.2695

Lupton, R., & Thornton, J. (2017). Disagreement, diversity, and participation: Examining the properties of several measures of political discussion network characteristics. *Political Behavior*, *39*, 585–608. https://doi.org/10.1007/s11109-016-9371-7

Luttrell, A., Philipp-Muller, A., & Petty, R. E. (2019). Challenging moral attitudes with moral messages. *Psychological Science*, *30*(8), 1136–1150. https://doi.org/10.1177/0956797619854706

Marchlewska, M., Górska, P., Green, R., Szczepańska, D., Rogoza, M., Molenda, Z., & Michalski, P. (2022). From individual anxiety to collective narcissism? Adult attachment styles and different types of national commitment. *Personality and Social Psychology Bulletin*, *50*(4), 495–515. https://doi.org/10.1177/01461672221139072

Marchlewska, M., Górska, P., Molenda, Z., Lipowska, K., & Malinowska, K. (2023). The fear of confession? High Catholic collective narcissism and low secure identification with Catholics predict increased pedophilia myth acceptance. *European Journal of Social Psychology*, *53*(2), 354–366. https://doi.org/10.1002/ejsp.2907

Marchlewska, M., Górska, P., Podsiadłowski, W., Rogoza, M., & Szczepańska, D. (2024). So different yet so alike? Political collective narcissism predicts blatant dehumanization of political out-groups among conservatives and liberals. *British Journal of Social Psychology*, *64*(2), Article 12803. https://doi.org/10.1111/bjso.12803

Martherus, J. L., Martinez, A. G., Piff, P. K., & Theodoridis, A. G. (2021). Party animals? Extreme partisan polarization and dehumanization. *Political Behavior*, *43*, 517–540.

Mason, L. (2016). A cross-cutting calm: How social sorting drives affective polarization. *Public Opinion Quarterly*, *80*(S1), 351–377. https://doi.org/10.1093/poq/nfw001

McCoy, J., & Somer, M. (2019). Toward a theory of pernicious polarization and how it harms democracies: Comparative evidence and possible remedies. *The Annals of the American Academy of Political and Social Science*, *681*(1), 234–271. https://doi.org/10.1177/000271621881878

Moore-Berg, S. L., Ankori-Karlinsky, L. O., Hameiri, B., & Bruneau, E. (2020). Exaggerated meta-perceptions predict intergroup hostility between American political partisans. *Proceedings of the National Academy of Sciences*, *117*(26), 14864–14872. https://doi.org/10.1073/pnas.2001263117

Nederhof, A. J. (1985). Methods of coping with social desirability bias: A review. *European Journal of Social Psychology*, *15*(3), 263–280. https://doi.org/10.1002/ejsp.2420150303

Page-Gould, E., Harris, K., MacInnis, C. C., Danyluck, C. M., & Miller, I. D. (2022). The intergroup perspective on cross-group friendship. *Advances in Experimental Social Psychology*, *65*, 1–56. https://doi.org/10.1016/bs.aesp.2021.10.001

Pettigrew, T. F. (1997). Generalized intergroup contact effects on prejudice. *Personality and Social Psychology Bulletin*, *23*(2), 173–185. https://doi.org/10.1177/0146167297232006

Pettigrew, T. F. (2009). Secondary transfer effect of contact: Do intergroup contact effects spread to noncontacted out-groups? *Social Psychology*, *40*(2), 55–65. https://doi.org/10.1027/1864-9335.40.2.55

Pettigrew, T. F., & Tropp, L. R. (2006). A meta-analytic test of intergroup contact theory. *Journal of Personality and Social Psychology*, *90*(5), 751–783. https://doi.org/10.1037/0022-3514.90.5.751

Pettigrew, T. F., & Tropp, L. R. (2008). How does intergroup contact reduce prejudice? Meta-analytic tests of three mediators. *European Journal of Social Psychology*, *38*(6), 922–934. https://doi.org/10.1002/ejsp.504

Pettigrew, T. F., Tropp, L. R., Wagner, U., & Christ, O. (2011). Recent advances in intergroup contact theory. *International Journal of Intercultural Relations*, *35*(3), 271–280. https://doi.org/10.1016/j.ijintrel.2011.03.001

Plichta, E. (2024). 2023 Polish parliamentary elections: Victory for women's rights or empty promises? *Femina Politica–Zeitschrift für feministische Politikwissenschaft*, *33*(1), 132–136. https://doi.org/10.3224/feminapolitica.v33i1.14

Robinson, C. (2010). Cross-cutting messages and political tolerance: An experiment using evangelical protestants. *Political Behavior*, *32*, 495–515. https://doi.org/10.1007/s11109-010-9118-9

Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, *48*, 1–36. https://doi.org/10.18637/jss.v048.i02

Rozin, P. (1999). The process of moralization. *Psychological Science*, *10*(3), 218–221. https://doi.org/10.1111/1467-9280.00139

Schein, C., & Gray, K. (2018). The theory of dyadic morality: Reinventing moral judgment by redefining harm. *Personality and Social Psychology Review*, *22*(1), 32–70. https://doi.org/10.1177/1088868317698288

Schmid, K., Hewstone, M., & Ramiah, A. A. (2013). Neighborhood diversity and social identity complexity: Implications for intergroup relations. *Social Psychological and Personality Science*, *4*(2), 135–142. https://doi.org/10.1177/1948550612446972

Simonsen, K. B., & Bonikowski, B. (2022). Moralizing immigration: Political framing, moral conviction, and polarization in the United States and Denmark. *Comparative Political Studies*, *55*(8), 1403–1436. https://doi.org/10.1177/00104140211060284

Sinclair, B. (2012). *The social citizen: Peer networks and political behavior*. University of Chicago Press.

Skitka, L. J. (2010). The psychology of moral conviction. *Social and Personality Psychology Compass*, *4*(4), 267–281. https://doi.org/10.1111/j.1751-9004.2010.00254.x

Skitka, L. J., Hanson, B. E., Morgan, G. S., & Wisneski, D. C. (2021). The psychology of moral conviction. *Annual Review of Psychology*, *72*(1), 347–366. https://doi.org/10.1146/annurev-psych-063020-030612

Skitka, L. J., Liu, J. H. F., Yang, Y., Chen, H., Liu, L., & Xu, L. (2013). Exploring the cross-cultural generalizability and scope of morally motivated intolerance. *Social Psychological and Personality Science*, *4*(3), 324–331. https://doi.org/10.1177/1948550612456404

Smith, C., Emerson, M., Gallagher, S., Kennedy, P., & Sikkink, D. (1999). *American Evangelicalism: Embattled and thriving*. University of Chicago Press.

Szczęśniak, A. (2024, April 5). Donald Tusk: "Możemy zmarnować zwycięstwo z 15 października. Idziemy łeb w łeb." *OKO.press*. https://oko.press/na-zywo/na-zywo-wybory-samorzadowe/tusk-zmarnowac-zwyciestwo

Tajfel, H. (1981). *Human groups and social categories*. Cambridge University Press.

Tausch, N., Birtel, M. D., Górska, P., Bode, S., & Rocha, C. (2024). A post-Brexit intergroup contact intervention reduces affective polarization between Leavers and Remainers short-term. *Communications Psychology*, *2*(1), Article 95. https://doi.org/10.1038/s44271-024-00146-w

van Zomeren, M., d'Amore, C., Pauls, I. L., Shuman, E., & Leal, A. (2024). The intergroup value protection model: A theoretically integrative and dynamic approach to intergroup conflict escalation in democratic societies. *Personality and Social Psychology Review*, *28*(2), 225–248. https://doi.org/10.1177/1088868323119212

van Zomeren, M., Postmes, T., & Spears, R. (2012). On conviction's collective consequences: Integrating moral conviction with the social identity model of collective action. *British Journal of Social Psychology*, *51*(1), 52–71. https://doi.org/10.1111/j.2044-8309.2010.02000.x

Verkuyten, M., Voci, A., & Pettigrew, T. F. (2022). Deprovincialization: Its importance for plural societies. *Social Issues and Policy Review*, *16*(1), 289–309. https://doi.org/10.1111/sipr.12082

Wagner, M. (2021). Affective polarization in multiparty systems. *Electoral Studies*, *69*, Article 102199. https://doi.org/10.1016/j.electstud.2020.102199

Wisneski, D. C., & Skitka, L. J. (2017). Moralization through moral shock: Exploring emotional antecedents to moral conviction. *Personality and Social Psychology Bulletin*, *43*(2), 139–150. https://doi.org/10.1177/0146167216676479

Wohl, M. J. A., Stefaniak, A., & Smeekes, A. (2020). Longing is in the memory of the beholder: Collective nostalgia content determines the method members will support to make their group great again. *Journal of Experimental Social Psychology*, *91*, Article 104044. https://doi.org/10.1016/j.jesp.2020.104044

Wojcieszak, M., & Warner, B. R. (2020). Can interparty contact reduce affective polarization? A systematic test of different forms of intergroup contact. *Political Communication*, *37*(6), 789–811. https://doi.org/10.1080/10584609.2020.1760406

Wölfer, R., & Hewstone, M. (2017). Beyond the dyadic perspective: 10 reasons for using social network analysis in intergroup contact research. *British Journal of Social Psychology*, *56*(3), 609–617. https://doi.org/10.1111/bjso.12195

Zaal, M. P., Saab, R., O'Brien, K., Jeffries, C., Barreto, M., & van Laar, C. (2017). You're either with us or against us! Moral conviction determines how the politicized distinguish friend from foe. *Group Processes & Intergroup Relations*, *20*(4), 519–539. https://doi.org/10.1177/1368430215615682

Zingora, T., Stark, T. H., & Flache, A. (2020). Who is most influential? Adolescents' intergroup attitudes and peer influence within a social network. *Group Processes & Intergroup Relations*, *23*(5), 684–709.

Zollinger, D. (2024). Place-based identities and cleavage formation in the knowledge society. *Electoral Studies*, *88*, Article 102768. https://doi.org/10.1111/ajps.12743

About the Authors

Paulina Górska is an assistant professor at the Faculty of Psychology, University of Warsaw, Poland. Her research interests involve collective action, intergroup contact, political polarization, and minority stress.

Dominika Bulska is a post-doc at the Faculty of Psychology, SWPS University, Poland, and a researcher at the Institute for Social Studies, University of Warsaw, Poland. In her research she focuses on prejudice, dehumanization, collective violence, and collective action.

Maciej Górski is a PhD student at the Faculty of Psychology, University of Warsaw, Poland, and a researcher at the Institute of Psychology, Polish Academy of Sciences, Poland. He is interested in cross-cultural differences in political ideology, prejudice, and well-being.

cogitatio

ARTICLE

Longitudinal Associations Between Perceived Inclusivity Norms and Opinion Polarization in Adolescence

Maor Shani [1,2] ⓘ, **Marjorie Berns** [1], **Lucy Bergen** [1], **Stefanie Richters** [1] ⓘ, **Kristina Krämer** [1], **Sophie de Lede** [1], and **Maarten van Zalk** [1] ⓘ

[1] Institute for Psychology, Osnabrück University, Germany
[2] Faculty of Social Sciences, Ariel University, Israel

Correspondence: Maor Shani (maor.shani@uni-osnabrueck.de)

Abstract

Affective polarization, characterized by emotional hostility and behavioral avoidance toward ideological opponents beyond mere policy disagreements, can pose a significant threat to social cohesion. However, this phenomenon remains relatively unexplored in adolescence. This longitudinal study investigates whether perceived inclusivity norms—emphasizing equality-based respect, open and constructive dialogue, and communal unity—reduce opinion-based affective polarization among adolescents. Using a sample of 839 students from two demographically distinct German secondary schools (grades 7–11), we developed and validated measures of polarization tailored to adolescents, capturing dialogue orientation and social distance toward ideological outgroups. Results revealed stable ideological subgroup differences in norms and attitudes, with conservative students exhibiting lower descriptive and prescriptive inclusivity norms and higher affective polarization compared to their liberal peers. However, a significant proportion of adolescents demonstrated fluid political orientations over time, highlighting the malleability of early political identities. Cross-lagged analyses showed no evidence that inclusivity norms directly reduce affective polarization, although early dialogue orientation significantly predicted greater social openness in diverse settings. Our findings advance the understanding of adolescent political identity development by demonstrating that while ideological orientations remain unstable during this period, group-based differences in norm perceptions and polarization tendencies are already evident. We emphasize how this developmental fluidity presents both opportunities and challenges for interventions, suggesting that effective depolarization strategies must account for the distinct characteristics of adolescent political socialization and the varying influence of school-based normative contexts.

1. Introduction

Affective polarization, characterized by heightened negative emotions and behavioral avoidance towards ideological opponents beyond policy disagreements (Bakker & Lelkes, 2024; Iyengar et al., 2019), is linked by substantial research to negative democratic consequences, including political gridlock, eroded trust, and the spread of mis/disinformation (Iyengar et al., 2019; Shah, 2025; Tang et al., 2022; Wagner, 2024). While research has extensively examined these detrimental outcomes and the development of such antagonistic orientations in adults, understanding their initial emergence during adolescence has received surprisingly limited attention (Boonen, 2015; Hooghe & Wilkenfeld, 2008; Hutchins, 2024; Lay et al., 2023; Rekker et al., 2015; Tyler & Iyengar, 2023; Wang et al., 2017). Adolescence is a critical period in political socialization, marked by identity formation and heightened social sensitivity (Dahl et al., 2018; Van Hoorn et al., 2017). During this stage, young individuals often develop politically charged identities and partisan-tinged hostilities, even lacking fully articulated partisan affiliations (Tyler & Iyengar, 2023). Recent evidence indicates affective polarization can manifest as early as age 11, making adolescence an essential window for understanding political animosities' roots (Hutchins, 2024).

Yet, significant methodological and theoretical gaps constrain our understanding of adolescent affective polarization, which this longitudinal study aims to address. First, traditional adult-focused polarization measures are often developmentally inappropriate for adolescents, whose political identities frequently center on specific contentious social issues rather than formal partisanship (Lay et al., 2023). We, therefore, developed and validated novel opinion-based polarization measures capturing how young people's issue-specific stances manifest in social distancing and diminished dialogue willingness with ideological opponents. Second, to address the dearth of developmental studies, we examined whether opinion-based groups exhibited differential patterns over time in openness to dialogue and social interaction, key behavioral indicators of affective polarization (Kekkonen et al., 2022). Third, while research has explored how personal beliefs drive intergroup hostility, fewer studies have systematically investigated the role of social norms, particularly within schools, in exacerbating or mitigating adolescent polarization (Bicchieri et al., 2023). Schools are vital socialization environments where adolescents internalize norms about intergroup behaviors via peer interactions and institutional messaging (González, 2024; Tankard & Paluck, 2016). Thus, we examined the potential of school-based inclusivity norms, perceptions that peers endorse and enact equality-based respect, open dialogue, and communal unity, to attenuate affective polarization.

1.1. Polarization and Its Affective Manifestation

In recent years, polarization has emerged as a fundamental challenge that is increasingly shaping political landscapes across the globe (Kingzette et al., 2021; McClosky, 1964; McCoy & Somer, 2019). While traditional polarization research focuses on ideological distance, *affective polarization* centers on the emotional dimension of intergroup relations in the political sphere (Bakker & Lelkes, 2024; Iyengar et al., 2019; Shah, 2025; Tang et al., 2022; Wagner, 2024). It involves negative feelings (dislike, distrust, anger) and

behavioral avoidance directed towards members of opposing political or opinion-based groups, independent of specific policy disagreements. This "us vs. them" mentality aligns with social identity theory, where group membership fosters ingroup loyalty and outgroup derogation (Bakker & Lelkes, 2024; Iyengar et al., 2012; Schedler, 2023).

Pertaining to democratic norms, a prominent hypothesis suggests that partisans who harbor more negative feelings toward opponents will be more inclined to overlook transgressions of democratic norms by representatives of their camp (Graham & Svolik, 2020; Svolik, 2019). As the tendency to dismiss or delegitimize opposing viewpoints increases, the capacity for inclusive governance and reasoned compromise deteriorates (Levendusky, 2013). Consistently, a robust correlation between affective polarization and democratic backsliding was found across 50 countries (Orhan, 2022). On the other hand, some research suggests that affective polarization might have potential positive effects, such as increasing voter turnout by heightening the perceived stakes of elections (Ferreira Da Silva & Garzia, 2024; Harteveld & Wagner, 2023). Another study demonstrated that across Germany, the UK, and the US, moderate levels of affective polarization strengthen democratic attitudes by increasing political engagement and representation, while both very low and extremely high levels of polarization are associated with weaker democratic support (Janssen & Turkenburg, 2024). Studies also show that citizens may often prioritize more relevant considerations over raw partisan sentiment when making political judgments (Broockman et al., 2023), and that "partisan animosity is not a unifying concept underpinning the psychology of polarization and democracy" (Voelkel et al., 2024, p. 9).

However, affective polarization may systematically erode institutional trust (Iyengar & Westwood, 2015; Reiljan & Ryan, 2021), impede constructive discourse, undermine societal cohesion by diminishing meaningful cross-group contact, and potentially catalyze societal divisions that culminate in discriminatory behaviors or intergroup conflict (Berntzen et al., 2024; Orian Harel et al., 2023). These group dynamics align with social identity theory's core propositions, which posit that group attachments can override rational policy deliberation (Greene, 2004; Tajfel & Turner, 2004). As identification with particular viewpoints solidifies, opposing groups transition from being perceived as fellow citizens to being viewed as antagonistic outgroups that threaten fundamental values (Stephan & Stephan, 2016), which also may infiltrate everyday social decisions, from employment preferences to interpersonal relationships (Gift & Gift, 2015; Huber & Malhotra, 2017). Understanding how these affective processes become embedded in group identities, and identifying strategies to mitigate resultant hostilities, is crucial for preserving democratic discourse and institutional legitimacy.

1.2. Affective Polarization in Adolescence

Recent research challenges traditional assumptions that political enmity emerges only in late adolescence (Greenstein, 1960; Hess & Torney-Purta, 2017), revealing that partisan-tinged hostility can emerge as early as age 11 (Kinder & Sears, 1985; Tyler & Iyengar, 2023). While affective polarization has been extensively studied in adults (e.g., Bakker & Lelkes, 2024; Iyengar et al., 2019; Kinder & Sears, 1985; Mason, 2015), its early manifestations remain comparatively understudied (Boonen, 2015; Hooghe & Wilkenfeld, 2008; Hutchins, 2024; Lay et al., 2023; Rekker et al., 2015; Tyler & Iyengar, 2023; Wang et al., 2017). Yet, these developmental years represent a critical window: children and adolescents often progress from general political awareness to emotionally charged ingroup–outgroup judgments (Hutchins, 2024; Lay et al., 2023). Far from being politically

neutral, many youths readily internalize antagonism toward opposing ideological groups, reflecting polarizing rhetoric encountered in their home environments and digital spaces (Iyengar et al., 2019; Shah, 2025; Van Deth et al., 2011).

Adolescence represents a period marked by heightened receptivity to social cues and identity-related norms (Dahl et al., 2018; Van Hoorn et al., 2017), with early-formed opinions and group memberships demonstrating remarkable stability across the lifespan (Hooghe & Wilkenfeld, 2008). Through interactions with parents, peers, teachers, and social media, adolescents internalize which viewpoints garner approval and which spark conflict (Boonen, 2019; Uzun & Lüküslü, 2023).

The consequences of early polarization can extend to political behavior: When youths enter adulthood with established polarized attitudes, their capacity for compromise and cross-ideological engagement may be compromised (Kingzette et al., 2021; Mason, 2018b). Research indicates that adolescents' negative orientations can match adults' partisan animosity in intensity (Boonen, 2019; Laffineur et al., 2024; Raabe & Beelmann, 2011), fostering patterns of outgroup distrust that transcend political domains. Therefore, cultivating dialogue across opinion-based group boundaries and developing critical civic skills during adolescence becomes essential for preventing entrenched polarization in adulthood (Reifen-Tagar & Cimpian, 2022).

1.3. Social Norms and Their Impact on Affective Polarization

A second crucial gap in the literature concerns the relatively limited exploration of how *perceived social norms* influence affective polarization. While research has extensively examined how personal beliefs drive negative partisanship, fewer studies have investigated how collective rules and standards, both explicit and implicit, can either reinforce or reduce intergroup hostility (Bicchieri et al., 2023; G. D. A. Brown et al., 2022; Crandall et al., 2002; Fehr et al., 2002; Paluck, 2009; Tankard & Paluck, 2016). Social norms function as informal regulatory mechanisms through two interrelated forms: prescriptive (injunctive) norms, which capture group-sanctioned appropriate behavior, and descriptive norms, which reflect perceptions of how group members actually behave (Bicchieri et al., 2023; Cialdini et al., 1990; Miller & Prentice, 2016; Smith et al., 2012). Within the *school* environment, normative influences on adolescent behavior become especially salient because they can inadvertently function as echo chambers that intensify us-them thinking (González, 2024; Nipedal et al., 2010; Váradi et al., 2021), yet they also hold the potential to foster respectful peer norms that reframe disagreement as constructive dialogue (González, 2024; Tropp et al., 2016). Importantly, perceived norms, even when based on misperceptions of peers' attitudes, can shape youth behavior toward either exclusion or tolerance (Váradi et al., 2021), regardless of actual group norms, capturing the standards that individuals recognize and internalize (Tankard & Paluck, 2016).

Ultimately, perceived norms present both challenges and opportunities for addressing polarization (González, 2024; Iyengar & Westwood, 2015; Mullinix & Lythgoe, 2023). While polarized ingroup norms can accelerate intergroup hostility (Motyl et al., 2014; Whitt et al., 2021), recalibrating group expectations may foster more open-minded attitudes, particularly during adolescence (Crandall et al., 2002). Educational institutions can establish institutional messages of fairness and respect, offering alternative frameworks for intergroup relations (Bracegirdle et al., 2022). In this context, perceived inclusivity norms (Schäfer et al., 2024) emerge as a significant form of prescriptive norm—when adolescents perceive peer support for equal

treatment, dialogue, and shared community membership, these norms may buffer against polarization by facilitating open discourse and reducing social distance from outgroup members. The following section examines how these inclusivity norms interact with youth socialization processes to potentially mitigate polarizing influences in educational settings.

1.4. Perceived Inclusivity Norms to Counter Affective Polarization

Inclusivity norms (Schäfer et al., 2024) comprise three interrelated components: (a) equality-based respect norms, (b) communal unity norms, and (c) constructive dialogue norms. Equality-based respect norms establish collective standards that recognize the inherent dignity of all group members, regardless of their ideological or background differences (Renger & Reese, 2017; Simon & Schaefer, 2016). Emphasizing such equal status enhances intergroup relations and reduces outgroup hostility (Simon et al., 2019). Communal unity norms, grounded in civic rather than ethnic conceptions of belonging, promote the understanding that groups can maintain distinct identities while participating in a shared public sphere (Verkuyten et al., 2020). Lastly, constructive dialogue norms foster respectful perspective-taking toward outgroup members, a process that reduces stereotyping and intergroup anxiety (Glasford & Dovidio, 2011; Gurin et al., 2013). While past research has examined equality-based respect, dialogue, and unity as individual beliefs, our focus lies on their potential when perceived as social norms. These norms represent a set of interrelated principles for managing differences in diverse societies, principles that can theoretically apply beyond cultural or ethnic diversity to ideological differences (Verkuyten et al., 2020). Together, these norms transcend mere passive acceptance by promoting active respect for ideological diversity (Schäfer et al., 2024). Such collective standards become particularly vital in contemporary contexts where political and opinion-based antipathy often remains unchallenged (Iyengar & Westwood, 2015; Mullinix & Lythgoe, 2023). These norms align with broader principles of intergroup tolerance (Gibson, 1992; Velthuis et al., 2021; Verkuyten, 2022; Verkuyten et al., 2019), fostering not just acceptance but active engagement in shared social spaces with those holding divergent beliefs. This framework encourages adolescents to view ideological differences as components of a diverse social fabric rather than zero-sum conflicts, potentially attenuating the emotional foundations of affective polarization (Levendusky & Stecula, 2023).

Initial research provides encouraging, albeit mixed, evidence for these norms' effectiveness. A recent multinational investigation revealed that perceived inclusivity norms correlate with enhanced tolerance and cooperation while reducing avoidance of ideological opponents, even among individuals who strongly disagree with opposing views (Schäfer et al., 2024). Additionally, peer-led equality-based respect programs in educational settings have demonstrated potential to foster positive outgroup attitudes, though with time-bounded effects (Shani et al., 2023). However, the relationship between inclusivity norms and polarization during adolescence has not been studied. Our longitudinal approach addresses this gap by examining potential bidirectional effects: whether inclusivity norms predict subsequent changes in affective polarization and whether affective polarization shapes the development of inclusivity norms.

1.5. The Current Study

This article examines how perceived inclusivity norms in educational settings might attenuate opinion-based polarization among adolescents. Focusing on early—to mid-adolescence, a formative period for opinion development, we address "opinion-based group" polarization rather than formal party identification (Lay et al., 2023). While partisan conflicts are visible manifestations, affective polarization also arises within

opinion-based groups around contentious issues like climate change or immigration (Bliuc et al., 2007; Mason, 2015, 2018a; McGarty et al., 2009). This approach aligns with the "policy-over-party" hypothesis, suggesting animus primarily reflects policy disagreements (Webster & Abramowitz, 2017). Though the competing "party-over-policy" perspective (polarization from social identity attachments to parties; Mason, 2018b) has gained traction, an issue-focus is particularly appropriate for adolescents still developing partisan identities and whose attitudes often center on specific issues. Research supports this: Even unbranded policy preferences can diminish partisanship's effect on interpersonal affect (Dias & Lelkes, 2022), suggesting policy positions may shape interpersonal evaluations more than partisan identities, especially for adolescents. In these contexts, individuals develop robust social identities around shared stances, fostering ingroup loyalty and hostility toward those with different views (Harteveld, 2021; Hobolt et al., 2021). Even without formal partisan identification, strongly held opinions can foster an us-them mentality, reinforcing societal divisions (Bakker & Lelkes, 2024).

Existing scholarship has not systematically examined how perceived inclusivity norms might reduce affective polarization in educational contexts, particularly for youth forming opinion-based rather than strictly partisan identities. While some interventions (e.g., Shani et al., 2023) show short-term tolerance gains, the processes of norm internalization and whether school demographic composition moderates such effects remain unclear. Moreover, conflicting findings on potential ideological asymmetries (Kekkonen et al., 2022; Kluge et al., 2024) highlight the need to investigate whether liberal- or conservative-leaning adolescents differ in receptivity to inclusivity norms.

To address these gaps, we conducted a two-wave longitudinal survey of students (grades 7–11) in two comprehensive secondary schools in Lower Saxony, Germany: one culturally diverse, one predominantly homogeneous. Embedded within the Together for Tolerance project (Shani et al., 2023), we analyzed data from these control schools between Time 1 (January-February 2023) and Time 2 (January 2024). This design enabled the examination of both prescriptive and descriptive inclusivity norms and their relationship to adolescent affective polarization; the intervention's effects on intergroup tolerance are reported elsewhere (Shani et al., 2023). Selecting two regional schools with distinct demographic profiles—one highly diverse (urban, high migration background) and one more homogeneous (less urban, predominantly German background)—allowed us to explore how the school norms-polarization relationship might differ contextually.

Germany provides a revealing context for studying opinion-based affective polarization among adolescents, as its multiparty system extends literature beyond typically studied two-party systems (e.g., the US). Adolescents may develop strong opinions on salient social issues (like LGBTQIA+ rights and refugee policies, prominent in German discourse; Schnetzer et al., 2024) even without stable party affiliations characteristic of two-party systems (Bantel, 2023; Schieferdecker et al., 2024; Wagner, 2021). Moreover, Germany's political landscape has transformed: traditional *Volksparteien* (mainstream parties) grow weaker while smaller parties gain prominence (Angenendt & Brause, 2024). This shift likely influences affective polarization, particularly as negative emotions may target emerging fringe parties. The success of parties like the far-right Alternative für Deutschland among young voters (Shukla & Otto, 2025) suggests heightened affective responses. Further, while Germany saw declining ideological polarization (1980–2010) with substantial policy consensus, affective polarization has steadily increased concurrently (Coffé et al., 2025; Ferreira Da Silva & Garzia, 2024). This divergence creates an ideal setting to examine how emotional divisions develop even when ideological differences are less stark (Helbling & Jungkunz, 2020).

Importantly, we do not empirically compare affective polarization or its relationship with inclusivity norms between adolescents and adults. Rather, we investigate how affective polarization across opinion-based groups emerges during adolescence and examine whether inclusivity norms can counteract such polarization by fostering engagement with opposing groups instead of withdrawal. Our comparative analysis is limited to interpreting our findings against existing research and theories on adolescent political development reviewed earlier.

Our investigation had three primary objectives. First, we developed a method to obtain opinion-based groups and examine affective polarization specifically adapted for adolescents. They often lack stable partisan identities, with affiliations centering on opinions about salient issues rather than formal partisanship, particularly in complex multi-party systems (Lay et al., 2023). To enhance ecological validity, participants first shared their opinions on LGBTQIA+ and refugee rights, issues highly relevant to German adolescents (Schnetzer et al., 2024). Participants then responded to vignette-based scenarios depicting peers with opposing stances, reporting their willingness to engage or distance themselves socially and politically. Specifically, social distance, individuals' willingness for social interaction with outgroup members (Bogardus, 1947), is a well-established affective polarization measure in political psychology (Druckman & Levendusky, 2019; Iyengar et al., 2019). While variously operationalized (e.g., general partisan-directed attitudes; Iyengar et al., 2012), our study adopted assessments of comfort with outgroup members across specific social contexts (e.g., as neighbors, in-laws) to capture concrete behavioral manifestations. Social distance measures are particularly useful for examining adolescent intergroup relations (e.g., Binder et al., 2009), as they assess tangible social preferences reflecting everyday political antipathy. Additionally, we included dialogue orientation measures (Cionea et al., 2014) to examine if affective polarization extends beyond social avoidance to impair democratic discourse, a critical concern, as animosity can bias information processing and lead to selective exposure (Kaiser et al., 2022; Lelkes & Westwood, 2017; Mosleh et al., 2025). This approach captured both interpersonal and civic dimensions of potential opinion-based hostility.

Second, we further classified students into opinion-based groups, based on their positions regarding LGBTQIA+ and refugee rights, we examined whether these opinion-based groups exhibited different patterns of perceived norms, affective polarization, and dialogue orientation over time. While research with adults consistently demonstrates ideological asymmetries—with conservatives often expressing stronger outgroup negativity (Iyengar et al., 2019; Kekkonen et al., 2022; Kluge et al., 2024; Peffley et al., 2024)—such patterns among adolescents remain unexplored. Understanding these early group-specific tendencies is crucial for identifying which individuals might be most vulnerable to, or protected from, polarizing influences in educational settings (Lay et al., 2023).

Finally, we investigated whether perceived inclusivity norms predicted reductions in affective polarization over time (Schäfer et al., 2024; Verkuyten et al., 2020). Using cross-lagged models, we assessed whether Time 1 norms predicted subsequent changes in dialogue orientation and social distance while examining two key moderating factors. First, we analyzed whether political ideology moderated norm receptivity, and second, whether school diversity influenced norm effectiveness. Heterogeneous environments may provide more conducive conditions for inclusivity messages (Mitchell, 2019), while homogeneous settings could attenuate norm salience. Additionally, we examined how demographic characteristics—including religious affiliation, migration background, and gender—moderated normative influences.

This comprehensive investigation advances our understanding of how opinion-based hostility emerges and can be moderated within adolescent communities, offering insights into both theoretical frameworks of polarization and empirically informed interventions during this critical developmental period.

2. Methods

2.1. Sample and Procedure

The study protocol received approval from both the authors' institutional ethics committee and the local school authority responsible for educational research. Students provided active opt-in consent before each data collection wave, while parents received opt-out consent forms. Participants completed the survey during regular class time using smartphones or tablets, with trained research personnel present to provide assistance and ensure confidentiality. The 45-minute survey assessed demographics, psychosocial variables, school activities, and social networks, with our analysis focusing on polarization-relevant measures (complete instrument available in the Supplementary File).

The initial sample comprised 860 students (329 from the diverse school, 531 from the nondiverse school) across two waves spaced one year apart. After excluding 21 participants who discontinued before page 20 (of 46) or exhibited completion times under 500 seconds, the final analytic sample included 839 students. The diverse school retained 317 participants (142 at Time 1 only, 26 at Time 2 only, 149 at both waves), while the nondiverse school maintained 522 participants (168 at Time 1 only, 75 at Time 2 only, 279 at both waves). A priori power analysis using the semPower package (Moshagen & Bader, 2023) indicated that 792 participants would provide adequate power (80%) to detect small cross-lagged effects in latent-variable CLPMs, confirming our sample size as sufficient.

Detailed demographics appear in the Supplementary File, Table S1. The sample comprised 45% female and 51% male participants (the remaining participants indicated other gender identities), with a mean age of 14.59 years ($SD = 1.47$). The schools demonstrated significant demographic differences: The diverse school reported higher proportions of students with migration backgrounds (72% vs. 34%, $p < .001$) and distinct religious compositions (37% Muslim, 33% Christian vs. 11% Muslim, 59% Christian in the nondiverse school, $p < .001$). Ethnocultural backgrounds varied substantially, with 60% of nondiverse school students identifying as of German origin, compared to more varied regional backgrounds in the diverse school (e.g., 24% Middle East/North Africa, 4.7% Sub-Saharan Africa). Political orientations (see Section 2.2) further differentiated the schools: The diverse school showed higher proportions of "anti-LGBTQIA+, pro-refugees" students (33% vs. 8.6%), while the nondiverse school had a greater representation of "Conservatives" (6.0% vs. 2.8%) and "pro-LGBTQIA+, anti-refugees" students (10% vs. 2.8%, $p < .001$).

2.2. Measures

All measures were administered in German using a translation/back-translation procedure conducted by the research team. All scale items are provided in the Supplementary File, Table S2, translated into English.

2.2.1. Perceived Prescriptive and Descriptive Inclusivity Norms

The prescriptive inclusivity norms scale comprised six items, with pairs of items assessing three core components: equality-based respect, unity, and dialogue. This measure was adapted from a previously validated 10-item scale (Shani et al., 2023). Participants evaluated their entire school community's agreement with each statement using a 5-point scale (1 = *do not agree at all*, 5 = *fully agree*). Equality-based respect items were adapted from established measures (Renger & Reese, 2017; Renger & Simon, 2011), while unity and dialogue items drew from the Interculturalism Index (Verkuyten et al., 2020). Exploratory factor analyses (EFA) using maximum likelihood estimation with Varimax rotation, conducted separately for each school and wave, consistently revealed a one-factor solution (see Supplementary File, Table S3). Factor retention was determined by eigenvalues greater than 1.0 and examination of explained variance proportions. High factor loadings supported construct unidimensionality, with the factor explaining 60–72% of total variance. Internal consistency was excellent across all subsets (Cronbach's α and McDonald's $\omega \geq .90$; see Supplementary File, Table S7).

It is important to note that while our prescriptive inclusivity norms measure uses examples of cultural, religious, and national diversity to contextualize the items for adolescent comprehension (as established during pretesting; see Supplementary File, Table S2, for specific items), the theoretical construct extends beyond ethnic/cultural inclusivity to encompass broader intergroup relations principles. Our factor analyses support the view that these items represent a coherent underlying construct rather than separate dimensions, suggesting participants interpreted them as reflecting general inclusivity principles rather than solely cultural diversity attitudes.

The descriptive inclusivity norms measure, developed through preliminary focus groups with six adolescents, consisted of four items. Two items assessed positive inclusivity behaviors and two captured exclusionary behaviors. Participants estimated the proportion of schoolmates engaging in each behavior using a 5-point scale (1 = *nobody*, 5 = *all*). EFA revealed a consistent two-factor structure distinguishing between positive and negative behaviors, explaining 33–57% of total variance across subsets (see Supplementary File, Table S4). Inter-item correlations were adequate for negative descriptive norms but demonstrated lower consistency for positive descriptive norms (see Supplementary File, Table S7).

2.2.2. Opinion-Based Polarization Assessment

Figure 1 presents a flowchart that describes the sequence of texts and questions included in the measure. Participants' assignment to a specific vignette profile (or a generic one) was based on their responses to the four yes/no questions about LGBTQIA+ rights (same-sex marriage, adoption) and refugee policies (admission levels, societal contribution). An option to skip the question was available.

Participants then encountered a vignette featuring a same-age, same-gender fictional peer (e.g., "Lisa is 14 years old") holding opposing views. Given that adolescents primarily orient themselves toward same-gender and same-age peers (McMillan, 2022; McPherson et al., 2001), this matching enhances the measure's ecological validity and relevance for adolescent respondents (Primi et al., 2016). The profile assignment logic aimed to present participants with a peer holding opposing views on at least one major issue domain, specifically:

Opinion-Based Polarization Assessment
Flowchart of Measurement Process

Step 1: Initial Opinion Questions
(Participants allowed to skip questions)

LGBTQIA+ Rights

Q1: Same-sex marriage

(For/Against)

Q2: LGBTQIA+ adoption

(For/Against)

Refugee Policies

Q3: Refugee admission

(For/Against)

Q4: Societal contribution

(Contribute/Worsen)

Step 2: Profile Assignment
(Presenting same-age, same gender peer with opposing views)

If pro-LGBTQIA+ and pro-refugee
→ Conservative profile

If anti-LGBTQIA+ and anti-refugee
→ Liberal profile

If pro-LGBTQIA+ and anti-refugee
→ Pro-refugee profile

If anti-LGBTQIA+ and pro-refugee
→ Pro-LGBTQIA+ profile

If inconsistent within either domain
(e.g., yes to marriage, no to adoption)
→ Generic profile: "someone whose
opinions are different than yours"

Step 3: Measures Following Profile Presentation

Dialogue Orientation

Three-item scale (1–5):

1. Attempt persuasion
2. Seek understanding
3. Explain position

Social Distance

Six-item scale (1–5):

Willingness to interact with peer

across contexts (classroom, home,

public settings)

Step 4: Opinion-Based Group Classification
(For analysis purposes)

Conservatives:
Anti-LGBTQIA+ AND Anti-refugee

Liberals:
Pro-LGBTQIA+ AND Pro-refugee

Hybrid Group 1:
Pro-LGBTQIA+ AND Anti-refugee

Hybrid Group 2:
Anti-LGBTQIA+ AND Pro-refugee

Inconsistent:
Inconsistent answers within
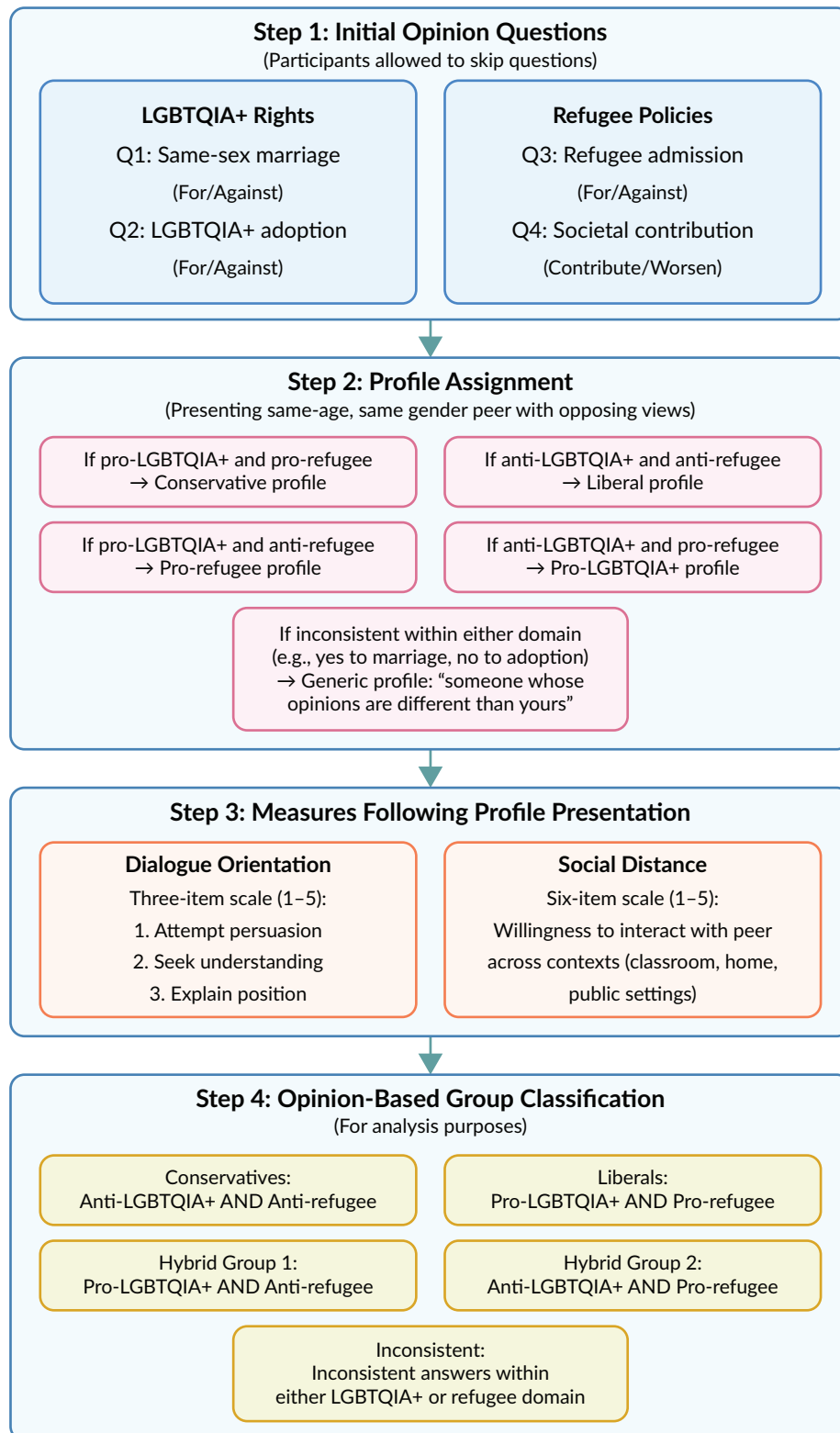either LGBTQIA+ or refugee domain

Figure 1. Flowchart depicting adolescent's opinion-based polarization assessment. Note: Classification enables comparison of dialogue orientation and social distance between groups.

1. Participants consistently (that is, matching *within*-domain answers) liberal on both domains (pro-LGBTQIA+, pro-refugee) received a consistently conservative profile (anti-LGBTQIA+, anti-refugee).
2. Participants consistently conservative on both domains (anti-LGBTQIA+, anti-refugee) received a consistently liberal profile (pro-LGBTQIA+, pro-refugee).
3. Participants liberal on LGBTQIA+ issues but conservative on refugee issues received a pro-refugee profile.
4. Participants conservative on LGBTQIA+ issues but liberal on refugee issues received a pro-LGBTQIA+ profile. Crucially, participants whose answers were inconsistent *within* either domain (e.g., answering "yes" to same-sex marriage but "no" to LGBTQIA+ adoption, or similar inconsistency on refugee questions), regardless of their stance on the other domain, were assigned the generic profile describing "someone whose opinions are different than yours." This group, termed "inconsistent" in the classification below, was thus treated separately from those holding clear, albeit mixed, stances across the two domains (e.g., consistently pro-LGBTQIA+ and consistently anti-refugee). In all conditions, the vignette was situated within participants' school environment.

The presentation of a profile was followed by two established measures:

1. Dialogue orientation: Adapted from Cionea et al.'s (2014) dialogue types, this three-item scale assessed willingness to (a) attempt persuasion, (b) seek understanding of others' perspectives, and (c) explain one's position (1 = *not at all*, 5 = *very much*). EFA using weighted least squares estimation confirmed a unidimensional structure (see Supplementary File, Table S5) with satisfactory reliability (see Supplementary File, Table S7).
2. Social distance (indicating affective polarization): A six-item scale measured willingness to interact with the hypothetical peer across contexts (e.g., classroom seating, home visits, public interactions) using a 5-point scale (1 = *very low*, 5 = *very high willingness*). EFA using minimum residual estimation with varimax rotation revealed a single-factor solution explaining 64–69% of variance (see Supplementary File, Table S6). Internal consistency was high ($\alpha > .90$) across schools and times (see Supplementary File, Table S7).

2.2.3. Classification of Opinion-Based Groups

Finally, to facilitate comparisons of norms and polarization measures between ideologically distinct subgroups, we categorized participants based on their pattern of responses to the four questions on LGBTQIA+ and refugee rights. Participants were classified into one of five groups:

1. Conservatives: Consistently anti-LGBTQIA+ (answered "no"/"against" to both items) AND consistently anti-refugee (answered "no"/"against" or "make things worse" to both items).
2. Liberals: Consistently pro-LGBTQIA+ (answered "yes"/"for" to both items) AND consistently pro-refugee (answered "yes"/"for" or "contribute" to both items). However, not all participants' stances aligned along a traditional left-right or liberal-conservative continuum. We classified these "hybrid groups" as:
3. Pro-LGBTQIA+, Anti-Refugees: Consistently pro-LGBTQIA+ AND consistently anti-refugee, and
4. Anti-LGBTQIA+, Pro-Refugees: Consistently anti-LGBTQIA+ AND consistently pro-refugee.
5. Inconsistent: Participants who provided inconsistent answers *within* either the LGBTQIA+ domain or the refugee domain (e.g., supporting same-sex marriage but opposing adoption, or vice-versa; similar

inconsistency on refugee items). This group captures individuals whose views did not align consistently even within a single-issue dimension.

2.3. Statistical Analysis

Initial data screening identified a programming error that resulted in incorrect profile assignments for 36 cases in the social distance and dialogue orientation measures; these cases were excluded from subsequent analyses. Missing data rates ranged from 1.7% to 6.9% across variables. Little's missing completely at random (MCAR) test yielded non-significant results ($p > .05$) for all school-wave subsets, supporting the MCAR assumption. Consequently, we employed robust analytical methods for handling missing data: mixed linear models and full information maximum likelihood (FIML) within structural equation modeling (SEM) frameworks. These approaches maximize available information without requiring listwise deletion, thereby enhancing estimate efficiency and statistical validity (Gabrio et al., 2022; Little & Rubin, 2020).

Preliminary analyses examined variable distributions (detailed in the Supplementary File, Tables S8–S12). Although Shapiro-Wilk tests indicated non-normality ($ps < .001$) and Henze-Zirkler tests revealed multivariate non-normality for both items ($HZ = 2.17$, $p < .01$) and scales ($HZ = 4.30$, $p < .01$; see Supplementary File, Tables S13–S14), all variables demonstrated acceptable skewness and kurtosis values (within ± 1). Given the established robustness of maximum likelihood estimation to moderate violations of normality assumptions (Enders, 2001; Kline & Little, 2023), we proceeded with the planned analyses.

Our analytical strategy comprised three main components. First, we assessed the temporal stability of opinion-based groups using a Sankey plot and marginal homogeneity tests. Second, we examined between-group differences using mixed linear models implemented in lme4 (Bates et al., 2015). These models accounted for the nested data structure (observations within participants) with fixed effects for wave, political orientation group, and their interaction, plus random intercepts for participants. Group comparisons utilized estimated marginal means (emmeans; Lenth, 2021), with Cohen's d effect sizes calculated from residual standard deviations. We applied Kenward-Roger adjustments for degrees of freedom and Tukey's method for multiple comparisons. Intraclass correlation coefficients quantified the stability of individual responses across waves.

Finally, we conducted longitudinal SEM to examine relationships among inclusivity norms, dialogue orientations, and social distance. After establishing measurement invariance (configural, metric, and partial scalar; see Meredith, 1993; Van De Schoot et al., 2012), we specified cross-lagged panel models incorporating stability paths, cross-lagged effects, and within-wave residual correlations. Examination of both directions of these longitudinal associations is important because, although prior literature suggests that perceived norms affect affective polarization, the alternative causal direction between these constructs has not been examined. Model fit evaluation followed Hu and Bentler's (1999) criteria (CFI $\geq$ .95, RMSEA $\leq$ .06, SRMR $\leq$ .08), with missing data handled through full information maximum likelihood estimation (Enders, 2001). To assess moderation effects, we conducted multi-group analyses comparing unconstrained and constrained models across opinion-based groups. Significant moderation effects ($\Delta\chi^2$, ΔCFI $\geq$ .01; Cheung & Rensvold, 2002) were then examined through systematic path comparisons and partial invariance testing.

3. Results

3.1. Distribution and Changes in Political Opinion-Based Groups

Analysis of political group distributions revealed substantial variability both within and across waves, with notably high proportions of inconsistent response patterns, meaning that they tended to both agree and disagree on different items within the same ideological domain. In the diverse school at Time 1, the distribution showed a relative balance between liberals (24.74%) and students holding anti-LGBTQIA+ but pro-refugee views (28.52%), with a smaller proportion of conservatives (2.41%). By Time 2, while the proportion of conservatives increased (8.57%), the percentage of students demonstrating inconsistent response patterns also rose (from 29.21% to 32%). The nondiverse school exhibited distinct patterns, with higher proportions of liberals (27.96% at Time 1) and a more substantial conservative presence (5.15% at Time 1, increasing to 9.89% at Time 2).

Longitudinal analysis of individual-level change, which included only participants who responded at both waves (diverse school: $n = 149$, nondiverse school: $n = 279$), revealed significant instability in political opinions, particularly in the diverse school ($\chi^2[4] = 18.097$, $p = .001$). The Sankey diagram (Figure 2) illustrates substantial flux in political orientations, with students demonstrating considerable movement between categories across waves. For instance, in the diverse school, of the 37 students initially classified as liberal, only 16 maintained this orientation at Time 2, with 11 shifting to inconsistent responses. Similarly, in the nondiverse school, while the overall distribution changes were marginally significant ($\chi^2[4] = 8.917$, $p = .063$), potentially due to small cell sizes, individual trajectories showed considerable volatility, with only 37 of 68 initial liberals maintaining their orientation.
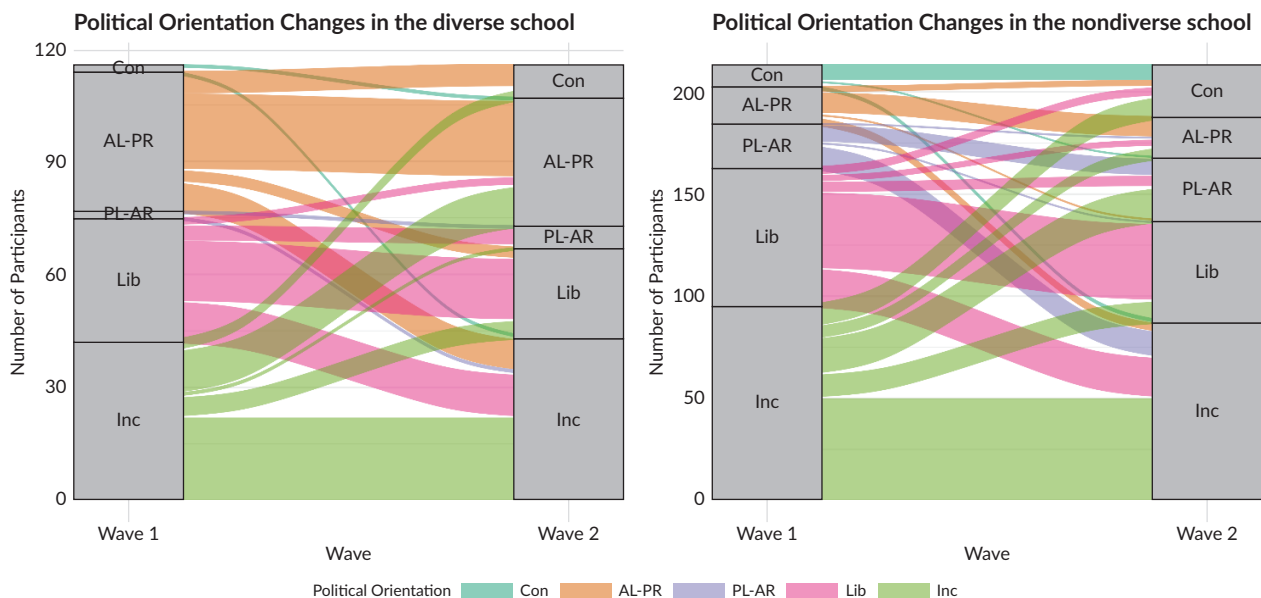

Figure 2. Political orientation (opinion-based groups) changes between Times 1 and 2 in the diverse and nondiverse schools. Note: Political orientations are categorized as conservative (Con), anti-LGBTQIA+/Pro-refugees (AL-PR), pro-LGBTQIA+/anti-refugees (PL-AR), liberal (Lib), and inconsistent (Inc).

3.2. Mixed Linear Models Predicting Outcomes From Political Groups and Time

Mixed linear models revealed significant main effects of political orientation group across all variables, with no significant time or interaction effects (Table S15 in the Supplementary File details model statistics; Table S16 presents estimated marginal means; see Tables S18–S22 for complete models). Group differences remained stable across waves, revealing several distinctive patterns.

Conservatives consistently demonstrated lower inclusivity-related attitudes compared to other groups. They reported significantly lower prescriptive norms than liberals at both waves (Time 1: $d = 0.57$, $p < .01$; Time 2: $d = 0.78$, $p < .001$) and lower than the anti-LGBTQIA+/pro-refugees group at Time 2 ($d = 0.93$, $p < .01$). They also perceived higher negative descriptive norms compared to liberals ($d = 0.68$, $p < .01$) and the anti-LGBTQIA+/pro-refugees group ($d = 0.78$, $p < .01$) at Time 1, with these differences persisting at Time 2. Conservatives exhibited the lowest dialogue readiness, particularly versus liberals (Time 1: $d = 1.00$, $p < .001$; Time 2: $d = 0.69$, $p < .01$) and the pro-LGBTQIA+/anti-refugees group (Time 1: $d = -0.78$, $p < .05$).

Liberals consistently exhibited more inclusive orientations across measures, reporting higher prescriptive norms than conservatives and higher positive descriptive norms at Time 1 ($d = 0.44$, $p < .05$). They maintained the highest dialogue readiness scores across waves, most notably compared to conservatives and the anti-LGBTQIA+/pro-Refugees group (Time 2: $d = 0.77$, $p < .001$).

The two hybrid groups exhibited distinct patterns. The anti-LGBTQIA+/pro-Refugees group reported higher prescriptive norms and lower negative descriptive norms than the pro-LGBTQIA+/anti-refugees group at Time 2 ($d = 0.56$ and $d = -0.65$ respectively, $ps < .05$), but demonstrated lower dialogue readiness. The pro-LGBTQIA+/anti-refugees group showed the lowest social distance scores, differing significantly from conservatives (Time 1: $d = 0.86$, $p < .05$; Time 2: $d = 0.48$, $p < .05$) and the anti-LGBTQIA+/pro-refugees group (Time 1: $d = 0.52$, $p < .05$).

3.3. Correlational Analysis

To provide a foundational overview of the relationships between key variables before examining longitudinal effects, we first examined bivariate Pearson correlations within and across waves, which are explained in the Supplementary File (Table S17).

3.4. Cross-Lagged SEM and Path Analysis

Initial analyses established measurement invariance across waves (see Supplementary File, Table S23). The measurement model demonstrated configural invariance (CFI = .961, RMSEA = .063) and metric invariance (ΔCFI = .000). Partial scalar invariance was achieved after releasing one item's intercept in the prescriptive norms scale ($\chi^2(197) = 613.57$, $p < .001$; CFI = .960; RMSEA = .060), supporting longitudinal comparisons of latent means.

The final structural model (Figure 3), incorporating theoretically justified residual covariances, demonstrated excellent fit ($\chi^2(368) = 679.91$, $p < .001$; CFI = .971; RMSEA = .032). All three constructs showed significant temporal stability from Time 1 to Time 2: prescriptive norms ($\beta = 0.299$), dialogue intentions ($\beta = 0.352$),
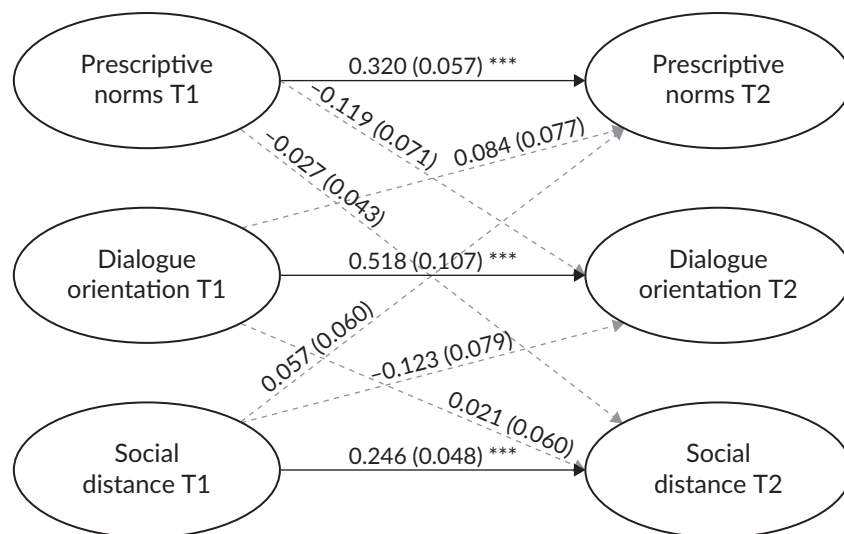
Figure 3. Structural paths in the modified cross-lagged panel model. Notes: Path coefficients are unstandardized. *SE* is the standard error of the unstandardized estimate and is presented in brackets; * $p < .05$, ** $p < .01$, *** $p < .001$.

and social distance ($\beta = 0.311$, all $ps < .001$). However, no significant cross-lagged effects emerged between constructs. Time 1 prescriptive norms neither predicted subsequent dialogue intentions ($\beta = -0.094$, $p = .093$) nor social distance ($\beta = 0.034$, $p = .521$), and initial levels of dialogue and social readiness did not predict later prescriptive norms (all $ps > .276$).

Moderation analyses by opinion-based groups employed observed variables due to convergence issues in the full SEM (for model fit indices see Supplementary File, Table S24). Given limited sample sizes for consistently classified participants across waves (conservative: $n = 8$; liberal: $n = 53$; anti-LGBTQIA+/pro-refugees: $n = 30$; pro-LGBTQIA+/anti-refugees: $n = 9$), group classification relied on Time 1 orientations ($n = 385$). Analyses revealed significant differences in temporal pathways ($\Delta\chi^2(27) = 42.392$, $p = .030$). A partial invariance model allowing dialogue orientation stability to vary across groups demonstrated good fit ($\chi^2(24) = 30.693$, $p = .163$; CFI = .948; RMSEA = .054, 90% CI [0.000, 0.104]). Dialogue orientation stability varied substantially: strongest for pro-LGBTQIA+/anti-refugees participants ($\beta = .524$, $p < .001$), followed by liberals ($\beta = .266$, $p < .001$) and anti-LGBTQIA+/pro-refugees participants ($\beta = .300$, $p = .022$), with a negative coefficient for conservatives ($\beta = -.204$, $p = .142$). All other pathways remained invariant.

School diversity analyses identified two significant path differences ($\chi^2(7) = 6.847$, $p = .445$; CFI = 1.000; RMSEA = 0.000, 90% CI [0.000, 0.060]). Social distance stability was stronger in the nondiverse school ($\beta = .330$, $p < .001$) compared to the diverse school ($\beta = .131$, $p = .093$), and higher Time 1 dialogue orientation predicted lower subsequent social distance only in the nondiverse school ($\beta = .204$, $p = .010$; diverse: $\beta = -.022$, $p = .696$).

4. Discussion

This study examined the relationship between perceived inclusivity norms in educational settings and opinion-based affective polarization during adolescence. Using a longitudinal design comparing two

schools with distinct multicultural compositions, we investigated whether perceived normative climate—independent of observed behavioral norms—could attenuate affective polarization over time. Despite theoretical propositions suggesting that normative emphasis on equality, respect, communal unity, and constructive dialogue would reduce polarization, our analyses revealed no evidence of cross-lagged causal effects from norms to either dialogue orientation or social distance. Below, we interpret these findings through intersecting developmental, contextual, and methodological perspectives.

4.1. Differences Across Political Orientations

Our mixed linear models revealed robust and consistent effects of political orientation on normative perceptions and intergroup engagement patterns, with ideological subgroups maintaining distinctive attitudinal profiles across waves. Conservative adolescents exhibited a consistent pattern characterized by lower perception of prescriptive norms, heightened perception of negative descriptive norms, and reduced propensity for cross-group dialogue. In contrast, liberal-identifying adolescents demonstrated the strongest inclusivity-oriented dispositions, consistently perceiving both prescriptive and positive descriptive norms while exhibiting enhanced dialogue readiness. Hybrid groups revealed more nuanced attitudinal configurations: the anti-LGBTQIA+/pro-refugees group perceived stronger prescriptive norms despite showing reduced dialogue readiness, while the pro-LGBTQIA+/anti-refugees group demonstrated notably lower social distance from ideological opponents.

These patterns align with research on ideological asymmetries showing conservative orientations associated with heightened outgroup distrust and reduced intergroup contact willingness (Brandt & Crawford, 2020; Ganzach & Schul, 2021). Conservative tendencies toward ingroup loyalty and resistance to social change (e.g., Feldman & Johnston, 2014; Jost, 2017) contrast with liberal characteristics of enhanced empathy and reduced ingroup emphasis (Jost et al., 2009; Morris, 2020). Our findings extend established work linking stronger ideological identities to increased political animosity (J. K. Brown & Hohman, 2022; Iyengar et al., 2019; Iyengar & Westwood, 2015; Mason, 2018a). Importantly, we found that many adolescents hold hybrid positions that transcend the traditional liberal-conservative divide. These mixed stances—such as supporting refugee policies while opposing LGBTQIA+ rights—can generate strong emotional attachments and issue-specific resistance to dialogue. While considerable instability was observed over one year, some stability patterns emerged, suggesting that for certain adolescents, ideological identities may crystallize earlier than previously recognized (Hess & Torney-Purta, 2017; Laffineur et al., 2024; Oosterhoff et al., 2021; Reifen-Tagar & Cimpian, 2022).

4.2. Methodological Insights and Developmental Patterns in Adolescent Political Identity Formation

Our vignette-based approach to measuring opinion-based affective polarization offered both methodological insights and theoretical revelations about adolescent political development. It presented concrete, developmentally appropriate scenarios rather than abstract political concepts; captured contextually relevant social issues; and measured both behavioral intentions and communicative orientations in naturalistic peer contexts. By focusing on concrete social scenarios rather than abstract political evaluations, our methodology captured behavioral manifestations of affective polarization through social distance (Druckman & Levendusky, 2019) and dialogue orientation measures (Cionea et al., 2014). However, the substantial proportion of participants showing within-individual fluctuations and

within-domain inconsistent responses revealed both methodological challenges and theoretically meaningful patterns in adolescent political identity formation.

The observed malleability in political orientations may reflect genuine opinions that are seemingly incoherent (i.e., hybrid on the conservative-liberal dichotomy) or inconsistent (i.e., mixed within domains), but may also be held by the wider public. Alternatively, it may reflect broader developmental processes, where youth actively explore and revise their beliefs through encounters with new information and social influences. This instability may stem from multiple factors: adolescents may still be developing the cognitive frameworks needed to organize political beliefs systematically, their opinions may be genuinely in flux as they engage with diverse perspectives, or they may lack sufficient exposure to political discourse to form stable viewpoints on complex social issues. Unlike adult populations, where ideological consistency often indicates political sophistication, adolescent hybrid groups (e.g., pro-LGBTQIA+, anti-refugees, or vice versa) and inconsistencies (e.g., supporting same-sex marriage but opposing adoption rights) may represent active engagement with complex political issues rather than measurement error.

Developmental perspectives provide theoretical grounds for expecting heightened instability and inconsistency during adolescence compared to adulthood. Political attitudes tend to become increasingly stable as individuals transition from adolescence to early adulthood (Rekker et al., 2019). Longitudinal research shows that core political attitudes in adulthood demonstrate remarkable stability over time, with approximately 75% of individuals maintaining consistent beliefs over decades (Peterson et al., 2020). The fluidity we observed aligns with developmental literature suggesting that adolescence represents a critical period for identity formation (Dahl et al., 2018), characterized by exploration across various domains, including political identity, coined by heightened susceptibility to social influence (Steinberg & Monahan, 2007), and ongoing cognitive development (Steinberg, 2005). Consequently, adolescent political attitudes might be particularly susceptible to change and less ideologically coherent.

However, while our findings highlight notable ideological inconsistencies and temporal instabilities in adolescents' political beliefs, we acknowledge that such patterns are not unique to adolescents. Research on European adults identified diverse belief structures, which were classified as "ideologue" groups (internally consistent left-right positions), "alternative" groups (multi-dimensional, often conflicting cultural and economic attitudes), and "unstructured" groups (weakly correlated beliefs; van Noord et al., 2025). Similarly, American adults frequently hold hybrid beliefs that combine economic liberalism with moral conservatism or vice versa (Baldassarri & Goldberg, 2014). Without a direct comparison of adolescents and adults using identical measures, we cannot conclude that the instability and ideological inconsistency observed in our sample exceeds that of adults. This limitation highlights the need for future research employing age-comparative designs with parallel measures to empirically test developmental differences in political attitude stability and consistency.

4.3. Inclusivity Norms and Affective Polarization

Contrary to our expectations that perceived inclusivity norms would predict increased dialogue and social orientations over time (Schäfer et al., 2024), cross-lagged analyses showed no longitudinal effects between inclusivity norms and readiness for cross-group social interactions. Instead, both dialogue orientation and social distance were primarily predicted by their own prior levels. This pattern suggests that naturally occurring

school inclusivity norms may be insufficient to overcome established identity-based attitudes. While previous research has demonstrated that school interventions can produce short-term tolerance gains (Shani et al., 2023), our findings indicate that the mere perception of inclusive norms does not alter dialogue orientation or social distance toward opposing groups. Family socialization may exert a stronger influence than school-based norms (Beck & Jennings, 1991; Sapiro, 2004; Tyler & Iyengar, 2023), or adolescents might respond more strongly to immediate peer-group norms than broader institutional climate messages. Future research should examine whether peer-led approaches prove more effective than institutional initiatives in shifting intergroup attitudes (Murrar et al., 2020).

Prescriptive inclusivity norms demonstrated stronger associations with polarization measures than descriptive norms, particularly negative descriptive norms, which showed no significant relationships across waves. This pattern aligns with Schäfer et al.'s (2024) findings of stronger connections between injunctive or prescriptive norms and tolerance across European countries. The limited influence of descriptive norms may reflect their context-specificity—prejudice often faces variable public condemnation depending on the target group (Crandall et al., 2002), suggesting that behavioral norms' effectiveness may vary between concrete outgroups and abstract ideological opponents.

4.4. Theoretical and Practical Implications

Our findings advance several key debates in political socialization and intergroup relations research. First, they support emerging perspectives that view political identity development as more fluid and complex than traditionally assumed. The substantial opinion-based group instability we observed challenges classic models suggesting political enmity emerges only in late adolescence (Greenstein, 1960; Hess & Torney-Purta, 2017). Rather, our results align with evidence that partisan hostility can emerge by age 11 (Tyler & Iyengar, 2023) while remaining malleable throughout adolescent development (Hooghe & Wilkenfeld, 2008; Lay et al., 2023).

Second, our findings illuminate the differential impact of norm types on adolescent intergroup attitudes. The stronger associations of affective polarization behavioral indicators with prescriptive norms compared to descriptive norms (see Supplementary File, Table S17) suggests that explicit institutional messages about inclusion may carry more weight than perceived peer behavior, supporting recent work on prescriptive norms as regulatory mechanisms in intergroup relations (Bicchieri et al., 2023; Cialdini & Jacobson, 2021). However, these normative influences compete with other socialization forces, particularly family influence in the transmission of partisan loyalties (Jennings & Niemi, 1968; Pedraza & Perry, 2020; Tyler & Iyengar, 2023). Our measure of inclusivity norms primarily contextualizes these principles within cultural diversity, which may influence their relationship with opinion-based polarization outcomes. Prejudice operates differently across ideological lines, with both liberals and conservatives expressing prejudice toward dissimilar outgroups (Crawford & Brandt, 2020). The modest associations we found between prescriptive inclusivity norms and opinion-based polarization may reflect this distinction between cultural and ideological inclusivity domains. Nevertheless, parallel research with adults using more generic inclusivity norm measures finds similar patterns of relationships between inclusivity norms and willingness to engage with ideological outgroups (Schäfer et al., 2024), suggesting some transferability of inclusivity principles across different intergroup contexts.

This raises the possibility of a bidirectional relationship between ideological orientation (rather than their affective consequences) and perceptions of inclusivity norms. Liberal-identifying adolescents demonstrated the strongest inclusivity-oriented dispositions, consistently perceiving positive prescriptive and descriptive norms while exhibiting enhanced dialogue readiness. If ideological groups differ in their approaches to various forms of diversity, with liberals often demonstrating greater endorsement of diversity-related attitudes (Crawford & Brandt, 2020), inclusivity norms may not represent a truly independent construct that influences polarization, but rather partially reflect existing ideological positions. The correlational nature of our study limits causal inferences, and the stability of both inclusivity norms and political positions over the one-year interval suggests a potential underlying relationship. Future research employing experimental manipulations of perceived norms could better disentangle these effects and test whether inclusivity norms can function as an intervention strategy that transcends ideological differences.

Third, the varying effects across contexts highlight the importance of considering multiple moderating factors. The enhanced dialogue-social openness link in nondiverse schools suggests that dialogue may be particularly crucial in settings where intergroup contact opportunities are limited. These findings bridge classic contact theory (Allport, 1954; Bracegirdle et al., 2022; Tausch et al., 2024) with contemporary research on how normative climates affect polarization in the absence of sustained outgroup interaction (Kingzette et al., 2021; Mullinix & Lythgoe, 2023). Although dialogue orientation was operationalized distinctly from perceived inclusivity norms, this result offers preliminary support for the theoretical role of inclusivity norms in reducing polarization, given that dialogue represents one of the core values associated with tolerance toward political adversaries (Schäfer et al., 2024).

These insights suggest several actionable interventions. Since adolescents often overestimate peer hostility, interventions that reveal more moderate actual peer attitudes can help recalibrate social perceptions (You & Lee, 2024). Face-to-face discussions combined with inclusivity messages can effectively reduce stereotypes and promote cross-group friendships (Levendusky & Stecula, 2021). Given the emotional foundations of polarization (Bakker & Lelkes, 2024), integrating inclusivity norms with emotion regulation strategies (Evers et al., 2014; Simonsson et al., 2022) may enhance intervention effectiveness. Programs should be tailored to ideological orientations—emphasizing communal values for conservative youth (Mullinix & Lythgoe, 2023) while adopting nuanced approaches with those holding hybrid beliefs to sustain dialogue without eliciting defensiveness.

4.5. Limitations and Future Directions

Several limitations qualify our conclusions. While our polarization measure enabled the assessment of opinion-based groups in an adolescent sample, its complexity and narrow focus on only two issues with two questions each resulted in a high proportion of within-domain inconsistent responses that complicated appropriate profiling and classification into opinion-based groups. Future research might develop more comprehensive profiles using questions that address a broader range of societal issues, or focus on partisan group affiliations rather than opinion-based groups.

Regarding our research design, the two-wave structure with a one-year interval limited our ability to detect potential short-term or delayed shifts in norms and polarization. Future studies employing longer timeframes with more frequent measurements (e.g., using random-intercept cross-lagged panel model,

Hamaker et al., 2015) could better illuminate whether and how norm-based changes emerge gradually. Finally, while we focused on perceived school norms, adolescents simultaneously navigate multiple social networks—including family, social media, and neighborhoods—that may transmit competing messages about outgroup engagement.

Attention should also be paid to the absence of a comparable adult sample. While our findings reveal considerable fluidity in adolescents' stances on LGBTQIA+ and refugee rights over a one-year period, without adult comparison using identical measures, we cannot definitively determine whether the instability observed is unique to adolescence or simply reflects broader patterns of opinion change across the lifespan. Future research should directly compare adolescent and adult samples using longitudinal designs and parallel measures to establish the relative stability and consistency of political attitudes across developmental stages. Moreover, instability and inconsistency may have resulted from our opinion-based polarization measure, which may have oversimplified adolescents' complex ideological explorations. Future studies should incorporate attitudinal ambivalence measures (Burger, 2024) and neutral response options to better distinguish between genuine political uncertainty, active ideological exploration, and measurement-related response inconsistency. Additionally, we did not assess strength of identification with opinion-based groups, though this factor often influences how ingroup norms are internalized or resisted (You & Lee, 2024).

The age range of our sample (grades 7–11) limits the generalizability to younger children or emerging adults, as political attitudes continue to develop across adolescence (Nieuwelink et al., 2018). Future studies should employ cohort-sequential designs that follow multiple age groups (including adolescents and adults of varying ages) over several years to disentangle age effects from cohort and period effects, and to identify the trajectories of opinion stability across the lifespan.

Furthermore, a key limitation pertains to our measurement of inclusivity norms. Following cognitive pretesting, we contextualized inclusivity norm items within cultural, religious, and national diversity to enhance comprehensibility for adolescent participants. This approach, while developmentally appropriate, may have narrowed the conceptual scope of inclusivity norms, potentially limiting their applicability to opinion-based polarization. Future research should develop and validate measures of inclusivity norms that more explicitly encompass ideological diversity alongside cultural diversity, allowing researchers to examine whether these norms function similarly or differently across domains.

Finally, while this study provides valuable insights into affective polarization among adolescents in Germany, several scope conditions should be considered when generalizing these findings to other contexts. Our measurement approach focusing on opinion-based rather than partisan polarization, while appropriate for adolescents with developing political identities in a multiparty system, might yield different results in contexts where partisan identities form earlier or where political conflict is structured differently. The divergence between Germany's declining ideological polarization and increasing affective polarization (Coffé et al., 2025; Ferreira Da Silva & Garzia, 2024) represents a unique political climate that may shape adolescent political socialization differently than in countries where ideological and affective polarization align. Additionally, Germany's historical legacies, including its Nazi past and East–West division and reunification, have created distinct political socialization patterns that influence how young people form and express political attitudes. The country's ongoing debates around LGBTQIA+ rights and refugee policies, which we used as focal issues, reflect particular national concerns that may manifest differently in countries

with distinct historical legacies and contemporary political narratives, yet these also vary across the federal states within Germany (Angenendt & Brause, 2024). Future research should explore these dynamics across diverse regional, national, cultural, and political settings to determine which aspects of adolescent affective polarization represent universal developmental processes versus context-specific manifestations shaped by particular political and educational environments.

These limitations suggest several promising research directions. The growing influence of algorithm-driven social media feeds and viral political content on adolescent attitudes may amplify or counteract tolerance-promoting messages in educational settings (Chan & Fu, 2017; Daniel, 2024), warranting investigation into how these digital environments interact with school-based norm interventions. Researchers should also examine the comparative effectiveness of peer-led initiatives versus institutional policies, explore how emotion regulation strategies might buffer against polarization-induced hostility, and investigate the dynamic interplay between evolving ideological alignments and normative influences during adolescent development.

Acknowledgments

Funding

Conflict of Interests

The authors declare no conflict of interests. In this article, editorial decisions were undertaken by Ulf R. Hedetoft (University of Copenhagen, Denmark).

Data Availability

Data and materials used in this study are available upon request to the corresponding author.

Supplementary Material

Supplementary material for this article is available online in the format provided by the author (unedited).

References

Allport, G. W. (1954). *The nature of prejudice*. Addison-Wesley.

Angenendt, M., & Brause, S. D. (2024). The long way towards polarized pluralism: Party and party system change in Germany. In T. Poguntke & W. Hofmeister (Eds.), *Political parties and the crisis of democracy* (1st ed., pp. 82–107). Oxford University Press. https://doi.org/10.1093/oso/9780198888734.003.0005 vBakker, B. N., & Lelkes, Y. (2024). Putting the affect into affective polarisation. *Cognition and Emotion*, *38*(4), 418–436. https://doi.org/10.1080/02699931.2024.2362366

Baldassarri, D., & Goldberg, A. (2014). Neither ideologues nor agnostics: Alternative voters' belief system in an age of partisan politics. *American Journal of Sociology*, *120*(1), 45–95. https://doi.org/10.1086/676042

Bantel, I. (2023). Camps, not just parties. The dynamic foundations of affective polarization in multi-party systems. *Electoral Studies*, *83*, Article 102614. https://doi.org/10.1016/j.electstud.2023.102614

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1). https://doi.org/10.18637/jss.v067.i01

Beck, P. A., & Jennings, M. K. (1991). Family traditions, political periods, and the development of partisan orientations. *The Journal of Politics*, *53*(3), 742–763. https://doi.org/10.2307/2131578

Berntzen, L. E., Kelsall, H., & Harteveld, E. (2024). Consequences of affective polarization: Avoidance, intolerance and support for violence in the United Kingdom and Norway. *European Journal of Political Research*, *63*(3), 927–949. https://doi.org/10.1111/1475-6765.12623

Bicchieri, C., Dimant, E., Gelfand, M., & Sonderegger, S. (2023). Social norms and behavior change: The interdisciplinary research frontier. *Journal of Economic Behavior & Organization*, *205*, A4–A7. https://doi.org/10.1016/j.jebo.2022.11.007

Binder, J., Zagefka, H., Brown, R., Funke, F., Kessler, T., Mummendey, A., Maquil, A., Demoulin, S., & Leyens, J.-P. (2009). Does contact reduce prejudice or does prejudice reduce contact? A longitudinal test of the contact hypothesis among majority and minority groups in three European countries. *Journal of Personality and Social Psychology*, *96*(4), 843–856. https://doi.org/10.1037/a0013470

Bliuc, A., McGarty, C., Reynolds, K., & Muntele, D. (2007). Opinion-based group membership as a predictor of commitment to political action. *European Journal of Social Psychology*, *37*(1), 19–32. https://doi.org/10.1002/ejsp.334

Bogardus, E. S. (1947). Measurement of personal-group relations. *Sociometry*, *10*(4), Article 306. https://doi.org/10.2307/2785570

Boonen, J. (2015). The development of stable party preferences: Explaining individual-level stability among adolescents in Belgium. *YOUNG*, *23*(4), 313–335. https://doi.org/10.1177/1103308815596887

Boonen, J. (2019). Learning who not to vote for: The role of parental socialization in the development of negative partisanship. *Electoral Studies*, *59*, 109–119. https://doi.org/10.1016/j.electstud.2019.04.001

Bracegirdle, C., Reimer, N. K., Van Zalk, M., Hewstone, M., & Wölfer, R. (2022). Disentangling contact and socialization effects on outgroup attitudes in diverse friendship networks. *Journal of Personality and Social Psychology*, *122*(1), 1–15. https://doi.org/10.1037/pspa0000240

Brandt, M. J., & Crawford, J. T. (2020). Worldview conflict and prejudice. In B. Gawronski (Ed.), *Advances in experimental social psychology* (Vol. 61, pp. 1–66). Elsevier. https://doi.org/10.1016/bs.aesp.2019.09.002

Broockman, D. E., Kalla, J. L., & Westwood, S. J. (2023). Does affective polarization undermine democratic norms or accountability? Maybe not. *American Journal of Political Science*, *67*(3), 808–828. https://doi.org/10.1111/ajps.12719

Brown, G. D. A., Lewandowsky, S., & Huang, Z. (2022). Social sampling and expressed attitudes: Authenticity preference and social extremeness aversion lead to social norm effects and polarization. *Psychological Review*, *129*(1), 18–48. https://doi.org/10.1037/rev0000342

Brown, J. K., & Hohman, Z. P. (2022). Extreme party animals: Effects of political identification and ideological extremity. *Journal of Applied Social Psychology*, *52*(5), 351–362. https://doi.org/10.1111/jasp.12863

Burger, A. M. (2024). Political ideology and attitudinal ambivalence: Investigating the role of ideological extremity. *Social Psychological and Personality Science*, *16*(4), Article 19485506231222958. https://doi.org/10.1177/19485506231222958

Chan, C., & Fu, K. (2017). The relationship between Cyberbalkanization and opinion polarization: Time-series analysis on Facebook pages and opinion polls during the Hong Kong occupy movement and the associated debate on political reform: Cyberbalkanization and opinion polarization. *Journal of Computer-Mediated Communication*, *22*(5), 266–283. https://doi.org/10.1111/jcc4.12192

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, *9*(2), 233–255. https://doi.org/10.1207/S15328007SEM0902_5

Cialdini, R. B., & Jacobson, R. P. (2021). Influences of social norms on climate change-related behaviors. *Current Opinion in Behavioral Sciences, 42*, 1–8. https://doi.org/10.1016/j.cobeha.2021.01.005

Cialdini, R. B., Reno, R. R., & Kallgren, C. A. (1990). A focus theory of normative conduct: Recycling the concept of norms to reduce littering in public places. *Journal of Personality and Social Psychology*, *58*(6), 1015–1026. https://doi.org/10.1037/0022-3514.58.6.1015

Cionea, I. A., Hample, D., & Fink, E. L. (2014). Dialogue types: A scale development study. In D. Mohammed & M. Lewiński (Eds.), *Virtues of argumentation: Proceedings of the 10th International Conference of the Ontario Society for the study of argumentation (OSSA)* (pp. 1–11). OSSA.

Coffé, H., Crawley, S., & Givens, J. (2025). Growing polarisation: Ideology and attitudes towards climate change. *West European Politics*. Advance online publication. https://doi.org/10.1080/01402382.2024.2435727

Crandall, C. S., Eshleman, A., & O'Brien, L. (2002). Social norms and the expression and suppression of prejudice: The struggle for internalization. *Journal of Personality and Social Psychology*, *82*(3), 359–378.

Crawford, J. T., & Brandt, M. J. (2020). Ideological (a)symmetries in prejudice and intergroup bias. *Current Opinion in Behavioral Sciences, 34*, 40–45. https://doi.org/10.1016/j.cobeha.2019.11.007

Dahl, R. E., Allen, N. B., Wilbrecht, L., & Suleiman, A. B. (2018). Importance of investing in adolescence from a developmental science perspective. *Nature*, *554*(7693), 441–450. https://doi.org/10.1038/nature25770

Daniel, J. (2024). Impact of social media usage on political polarization in the United States. *American Journal of Communication*, *6*(1), 56–66. https://doi.org/10.47672/ajc.1932

Dias, N., & Lelkes, Y. (2022). The nature of affective polarization: Disentangling policy disagreement from partisan identity. *American Journal of Political Science*, *66*(3), 775–790.

Druckman, J. N., & Levendusky, M. S. (2019). What do we measure when we measure affective polarization? *Public Opinion Quarterly*, *83*(1), 114–122. https://doi.org/10.1093/poq/nfz003

Eckstein, K., & Noack, P. (2015). Classroom climate effects on adolescents' orientations toward political behaviors. In P. Thijssen, J. Siongers, J. Van Laer, J. Haers, & S. Mels (Eds.), *Political engagement of the young in Europe: Youth in the crucible* (pp. 161–177). Routledge.

Enders, C. K. (2001). The impact of nonnormality on full information maximum-likelihood estimation for structural equation models with missing data. *Psychological Methods*, *6*(4), 352–370. https://doi.org/10.1037/1082-989X.6.4.352

Evers, C., Hopp, H., Gross, J. J., Fischer, A. H., Manstead, A. S. R., & Mauss, I. B. (2014). Emotion response coherence: A dual-process perspective. *Biological Psychology*, *98*, 43–49. https://doi.org/10.1016/j.biopsycho.2013.11.003

Fehr, E., Fischbacher, U., & Gächter, S. (2002). Strong reciprocity, human cooperation, and the enforcement of social norms. *Human Nature*, *13*(1), 1–25. https://doi.org/10.1007/s12110-002-1012-7

Feldman, S., & Johnston, C. (2014). Understanding the determinants of political ideology: Implications of structural complexity. *Political Psychology*, *35*(3), 337–358. https://doi.org/10.1111/pops.12055

Ferreira Da Silva, F., & Garzia, D. (2024). Affective polarization towards parties and leaders, and electoral participation in 13 parliamentary democracies, 1980–2019. *Public Opinion Quarterly*, *88*(4), 1234–1248. https://doi.org/10.1093/poq/nfae053

Gabrio, A., Plumpton, C., Banerjee, S., & Leurent, B. (2022). Linear mixed models to handle missing at random data in trial-based economic evaluations. *Health Economics*, *31*(6), 1276–1287. https://doi.org/10.1002/hec.4510

Ganzach, Y., & Schul, Y. (2021). Partisan ideological attitudes: Liberals are tolerant; the intelligent are intolerant. *Journal of Personality and Social Psychology*, *120*(6), 1551–1566. https://doi.org/10.1037/pspi0000324

Gibson, J. L. (1992). Alternative measures of political tolerance: Must tolerance be "least-liked"? *American Journal of Political Science*, *36*(2), 560. https://doi.org/10.2307/2111491

Gift, K., & Gift, T. (2015). Erratum to: Does politics influence hiring? Evidence from a randomized experiment. *Political Behavior*, *37*(3), 677–677. https://doi.org/10.1007/s11109-015-9314-8

Glasford, D. E., & Dovidio, J. F. (2011). E pluribus unum: Dual identity and minority group members' motivation to engage in contact, as well as social change. *Journal of Experimental Social Psychology*, *47*(5), 1021–1024. https://doi.org/10.1016/j.jesp.2011.03.021

González, R. (2024). The role of norms as promoters of social change. *Political Psychology*, *46*(2), 455–481. https://doi.org/10.1111/pops.13025

Graham, M. H., & Svolik, M. W. (2020). Democracy in America? Partisanship, polarization, and the robustness of support for democracy in the United States. *American Political Science Review*, *114*(2), 392–409. https://doi.org/10.1017/S0003055420000052

Greene, S. (2004). Social identity theory and party identification[*]. *Social Science Quarterly*, *85*(1), 136–153. https://doi.org/10.1111/j.0038-4941.2004.08501010.x

Greenstein, F. I. (1960). The benevolent leader: Children's images of political authority. *American Political Science Review*, *54*(4), 934–943. https://doi.org/10.2307/1952644

Gurin, P., Nagda, B. A., & Zúñiga, X. (2013). *Dialogue across difference: Practice, theory, and research on intergroup dialogue*. Russell Sage Foundation.

Hamaker, E. L., Kuiper, R. M., & Grasman, R. P. P. P. (2015). A critique of the cross-lagged panel model. *Psychological Methods*, *20*(1), 102–116. https://doi.org/10.1037/a0038889

Harteveld, E. (2021). Fragmented foes: Affective polarization in the multiparty context of the Netherlands. *Electoral Studies*, *71*, Article 102332. https://doi.org/10.1016/j.electstud.2021.102332

Harteveld, E., & Wagner, M. (2023). Does affective polarisation increase turnout? Evidence from Germany, The Netherlands and Spain. *West European Politics*, *46*(4), 732–759. https://doi.org/10.1080/01402382.2022.2087395

Helbling, M., & Jungkunz, S. (2020). Social divides in the age of globalization. *West European Politics*, *43*(6), 1187–1210.

Hess, R. D., & Torney-Purta, J. (2017). *The development of political attitudes in children*. Routledge.

Hobolt, S. B., Leeper, T. J., & Tilley, J. (2021). Divided by the vote: Affective polarization in the wake of the Brexit referendum. *British Journal of Political Science*, *51*(4), 1476–1493. https://doi.org/10.1017/S0007123420000125

Hooghe, M., & Wilkenfeld, B. (2008). The stability of political attitudes and behaviors across adolescence and early adulthood: A comparison of survey data on adolescents and young adults in eight countries. *Journal of Youth and Adolescence*, *37*(2), 155–167. https://doi.org/10.1007/s10964-007-9199-x

Hu, L.-t., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Huber, G. A., & Malhotra, N. (2017). Political homophily in social relationships: Evidence from online dating behavior. *The Journal of Politics*, *79*(1), 269–283. https://doi.org/10.1086/687533

Hutchins, R. (2024). The political polarization of pre-adolescents: Affective polarization in America's youth. *Social Education*, *88*(1), 20–23. https://doi.org/10.3102/00346543241288245

Iyengar, S., Lelkes, Y., Levendusky, M., Malhotra, N., & Westwood, S. J. (2019). The origins and consequences

of affective polarization in the United States. *Annual Review of Political Science*, *22*(1), 129–146. https://doi.org/10.1146/annurev-polisci-051117-073034

Iyengar, S., Sood, G., & Lelkes, Y. (2012). Affect, not ideology. *Public Opinion Quarterly*, *76*(3), 405–431. https://doi.org/10.1093/poq/nfs038

Iyengar, S., & Westwood, S. J. (2015). Fear and loathing across party lines: New evidence on group polarization. *American Journal of Political Science*, *59*(3), 690–707. https://doi.org/10.1111/ajps.12152

Janssen, L., & Turkenburg, E. (2024). Breaking free from linear assumptions: Unravelling the relationship between affective polarization and democratic support. *European Journal of Political Research*. Advance online publication. https://doi.org/10.1111/1475-6765.12725

Jennings, M. K., & Niemi, R. G. (1968). The transmission of political values from parent to child. *American Political Science Review*, *62*(1), 169–184.

Jost, J. T. (2017). Ideological asymmetries and the essence of political psychology. *Political Psychology*, *38*(2), 167–208. https://doi.org/10.1111/pops.12407

Jost, J. T., Federico, C. M., & Napier, J. L. (2009). Political ideology: Its structure, functions, and elective affinities. *Annual Review of Psychology*, *60*(1), 307–337. https://doi.org/10.1146/annurev.psych.60.110707.163600

Kaiser, J., Vaccari, C., & Chadwick, A. (2022). Partisan blocking: Biased responses to shared misinformation contribute to network polarization on social media. *Journal of Communication*, *72*(2), 214–240. https://doi.org/10.1093/joc/jqac002

Kekkonen, A., Suuronen, A., Kawecki, D., & Strandberg, K. (2022). Puzzles in affective polarization research: Party attitudes, partisan social distance, and multiple party identification. *Frontiers in Political Science*, *4*, Article 920567. https://doi.org/10.3389/fpos.2022.920567

Kinder, D. R., & Sears, D. O. (1985). Public opinion and political action. In G. Lindzey & E. Aronson (Eds.), *The handbook of social psychology* (3rd ed., pp. 659–741). Random House.

Kingzette, J., Druckman, J. N., Klar, S., Krupnikov, Y., Levendusky, M., & Ryan, J. B. (2021). How affective polarization undermines support for democratic norms. *Public Opinion Quarterly*, *85*(2), 663–677. https://doi.org/10.1093/poq/nfab029

Kline, R. B., & Little, T. D. (2023). *Principles and practice of structural equation modeling* (5th ed.). The Guilford Press.

Kluge, A., Adler, E., Nir, L., Halperin, E., Sams, M., & Levy, J. (2024). Asymmetry in political polarization at multiple levels of bias. *Political Psychology*, *45*(6), 1105–1121. https://doi.org/10.1111/pops.12967

Laffineur, C., Kasper, J., Brinke, L. T., Schumacher, G., & Bakker, B. N. (2024). *Political activism, interest, and affective polarization: Dutch adolescents and adults are not that different*. PsyArXiv. https://doi.org/10.31234/osf.io/aux5h

Lay, J. C., Holman, M. R., Greenlee, J. S., Oxley, Z. M., & Bos, A. L. (2023). Partisanship on the playground: Expressive party politics among children. *Political Research Quarterly*, *76*(3), 1249–1264. https://doi.org/10.1177/10659129221132223

Lelkes, Y., & Westwood, S. J. (2017). The limits of partisan prejudice. *The Journal of Politics*, *79*(2), 485–501. https://doi.org/10.1086/688223

Lenth, R. V. (2021). *emmeans: Estimated marginal means, aka least-squares means (R package version 1.7.0)*. https://CRAN.R-project.org/package=emmeans

Levendusky, M. S. (2013). Why do partisan media polarize viewers? *American Journal of Political Science*, *57*(3), 611–623. https://doi.org/10.1111/ajps.12008

Levendusky, M. S., & Stecula, D. A. (2021). *We need to talk: How cross-party dialogue reduces affective polarization* (1st ed.). Cambridge University Press. https://doi.org/10.1017/9781009042192

Little, R. J. A., & Rubin, D. B. (2020). *Statistical analysis with missing data* (3rd ed.). Wiley. https://doi.org/10.1002/9781119482260

Mason, L. (2015). "I disrespectfully agree": The differential effects of partisan sorting on social and issue polarization. *American Journal of Political Science*, *59*(1), 128–145. https://doi.org/10.1111/ajps.12089

Mason, L. (2018a). Ideologues without issues: The polarizing consequences of ideological identities. *Public Opinion Quarterly*, *82*(S1), 866–887. https://doi.org/10.1093/poq/nfy005

Mason, L. (2018b). *Uncivil agreement: How politics became our identity*. The University of Chicago Press.

McClosky, H. (1964). Consensus and ideology in American politics. *American Political Science Review*, *58*(2), 361–382. https://doi.org/10.2307/1952868

McCoy, J., & Somer, M. (2019). Toward a theory of pernicious polarization and how it harms democracies: Comparative evidence and possible remedies. *The ANNALS of the American Academy of Political and Social Science*, *681*(1), 234–271. https://doi.org/10.1177/0002716218818782

McGarty, C., Bliuc, A., Thomas, E. F., & Bongiorno, R. (2009). Collective action as the material expression of opinion-based group membership. *Journal of Social Issues*, *65*(4), 839–857. https://doi.org/10.1111/j.1540-4560.2009.01627.x

McMillan, C. (2022). Strong and weak tie homophily in adolescent friendship networks: An analysis of same-race and same-gender ties. *Network Science*, *10*(3), 283–300. https://doi.org/10.1017/nws.2022.24

McPherson, M., Smith-Lovin, L., & Cook, J. M. (2001). Birds of a feather: Homophily in social networks. *Annual Review of Sociology*, *27*(1), 415–444. https://doi.org/10.1146/annurev.soc.27.1.415

Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika*, *58*(4), 525–543. https://doi.org/10.1007/BF02294825

Miller, D. T., & Prentice, D. A. (2016). Changing norms to change behavior. *Annual Review of Psychology*, *67*(1), 339–361. https://doi.org/10.1146/annurev-psych-010814-015013

Mitchell, J. (2019). Prejudice in the classroom: A longitudinal analysis of anti-immigrant attitudes. *Ethnic and Racial Studies*, *42*(9), 1514–1533. https://doi.org/10.1080/01419870.2018.1493209

Morris, S. G. (2020). Empathy and the liberal-conservative political divide in the U.S. *Journal of Social and Political Psychology*, *8*(1), 8–24. https://doi.org/10.5964/jspp.v8i1.1102

Moshagen, M., & Bader, M. (2023). semPower: General power analysis for structural equation models. *Behavior Research Methods*, *56*(4), 2901–2922. https://doi.org/10.3758/s13428-023-02254-7

Mosleh, M., Martel, C., & Rand, D. G. (2025). Psychological underpinnings of partisan bias in tie formation on social media. *Journal of Experimental Psychology: General*, *154*(2), 378–390. https://doi.org/10.1037/xge0001662

Motyl, M., Iyer, R., Oishi, S., Trawalter, S., & Nosek, B. A. (2014). How ideological migration geographically segregates groups. *Journal of Experimental Social Psychology*, *51*, 1–14. https://doi.org/10.1016/j.jesp.2013.10.010

Mullinix, K. J., & Lythgoe, T. (2023). Priming norms to combat affective polarization. *Political Research Quarterly*, *76*(1), 186–199. https://doi.org/10.1177/10659129211073319

Murrar, S., Campbell, M. R., & Brauer, M. (2020). Exposure to peers' pro-diversity attitudes increases inclusion and reduces the achievement gap. *Nature Human Behaviour*, *4*(9), 889–897. https://doi.org/10.1038/s41562-020-0899-5

Nieuwelink, H., Ten Dam, G., Geijsel, F., & Dekker, P. (2018). Growing into politics? The development of adolescents' views on democracy over time. *Politics*, *38*(4), 395–410.

Nipedal, C., Nesdale, D., & Killen, M. (2010). Social group norms, school norms, and children's aggressive intentions. *Aggressive Behavior*, *36*(3), 195–204. https://doi.org/10.1002/ab.20342

Oosterhoff, B., Poppler, A., & Palmer, C. A. (2021). *Early adolescent political tribalism: Peer network homophily in political attitudes and values*. PsyArXiv. https://doi.org/10.31234/osf.io/8pfkm

Orhan, Y. E. (2022). The relationship between affective polarization and democratic backsliding: Comparative evidence. *Democratization*, *29*(4), 714–735. https://doi.org/10.1080/13510347.2021.2008912

Orian Harel, T., Shuman, E., Maoz, I., Balmas, M., & Halperin, E. (2023). Examining real-world legitimization of cross-party violence through two explanatory frameworks: Affective polarization and low group efficacy. *Journal of Social and Political Psychology*, *11*(1), 309–329. https://doi.org/10.5964/jspp.10007

Paluck, E. L. (2009). Reducing intergroup prejudice and conflict using the media: A field experiment in Rwanda. *Journal of Personality and Social Psychology*, *96*(3), 574–587. https://doi.org/10.1037/a0011989

Pedraza, F. I., & Perry, B. N. (2020). Validating a measure of perceived parent–child political socialization. *Political Research Quarterly*, *73*(3), 623–637. https://doi.org/10.1177/1065912919850632

Peffley, M., Yair, O., & Hutchison, M. L. (2024). Left-right social identity and the polarization of political tolerance. *Political Research Quarterly*, *77*(1), 30–44. https://doi.org/10.1177/10659129231189759

Peterson, J. C., Smith, K. B., & Hibbing, J. R. (2020). Do people really become more conservative as they age? *The Journal of Politics*, *82*(2), 600–611. https://doi.org/10.1086/706889

Primi, R., Zanon, C., Santos, D., De Fruyt, F., & John, O. P. (2016). Anchoring vignettes: Can they make adolescent self-reports of social-emotional skills more reliable, discriminant, and criterion-valid? *European Journal of Psychological Assessment*, *32*(1), 39–51. https://doi.org/10.1027/1015-5759/a000336

Raabe, T., & Beelmann, A. (2011). Development of ethnic, racial, and national prejudice in childhood and adolescence: A multinational meta-analysis of age differences: Age differences in prejudice. *Child Development*, *82*(6), 1715–1737. https://doi.org/10.1111/j.1467-8624.2011.01668.x

Reifen-Tagar, M., & Cimpian, A. (2022). Political ideology in early childhood: Making the case for studying young children in political psychology. *Political Psychology*, *43*(S1), 77–105. https://doi.org/10.1111/pops.12853

Reiljan, A., & Ryan, A. (2021). Ideological tripolarization, partisan tribalism and institutional trust: The foundations of affective polarization in the Swedish multiparty system. *Scandinavian Political Studies*, *44*(2), 195–219. https://doi.org/10.1111/1467-9477.12194

Rekker, R., Keijsers, L., Branje, S., & Meeus, W. (2015). Political attitudes in adolescence and emerging adulthood: Developmental changes in mean level, polarization, rank-order stability, and correlates. *Journal of Adolescence*, *41*(1), 136–147. https://doi.org/10.1016/j.adolescence.2015.03.011

Rekker, R., Keijsers, L., Branje, S., & Meeus, W. (2019). The formation of party preference in adolescence and early adulthood: How and when does it occur in the multiparty context of The Netherlands? *YOUNG*, *27*(1), 48–68. https://doi.org/10.1177/1103308818757037

Renger, D., & Reese, G. (2017). From equality-based respect to environmental activism: Antecedents and consequences of global identity. *Political Psychology*, *38*(5), 867–879. https://doi.org/10.1111/pops.12382

Renger, D., & Simon, B. (2011). Social recognition as an equal: The role of equality-based respect in group life. *European Journal of Social Psychology*, *41*(4), 501–507. https://doi.org/10.1002/ejsp.814

Sapiro, V. (2004). Not you parents' political socialization: Introduction for a new generation. *Annual Review of Political Science*, *7*(1), 1–23. https://doi.org/10.1146/annurev.polisci.7.012003.104840

Schäfer, L. F., Tausch, N., Bukowski, M., Jaspers, E., Lubbers, M. J., van Zalk, M., Ciordia, A., Potoczek, A., Estevan-Reina, L., Shani, M., Simons, J.-W., Friehs, M.-T., Gurbisz, D., Middendorf, W. M., Schäfer, S. J., Ullenboom, J., Graf, S., Hjerm, M., Lavest, C., . . . Christ, O. (2024). Inclusivity norms counter the negative societal effects of opinion polarization across 12 European countries. *PsyArXiv*. https://doi.org/10.31234/osf.io/9gh6k

Schedler, A. (2023). Rethinking political polarization. *Political Science Quarterly*, *138*(3), 335–359. https://doi.org/10.1093/psquar/qqad038

Schieferdecker, D., Joly, P., & Faas, T. (2024). Affective polarization between opinion-based groups in a context of low partisan discord: Measuring its prevalence and consequences. *International Journal of Public Opinion Research*, *36*(2). https://doi.org/10.1093/ijpor/edae009

Schnetzer, S., Hampel, K., & Hurrelmann, K. (2024). *Trendstudie Jugend in Deutschland-Verantwortung für die Zukunft? Ja, aber*. Datajockey Verlag. http://doi.org/10.17879/58918451646

Shah, T. M. (2025). Trends in political science research: Affective polarization. *International Political Science Abstracts*, *75*(1), 1–14. https://doi.org/10.1177/00208345251323534

Shani, M., de Lede, S., Richters, S., Kleuker, M., Middendorf, W., Liedtke, J., Witolla, S., & van Zalk, M. (2023). A social network intervention to improve adolescents' intergroup tolerance via norms of equality-based respect: The "together for tolerance" feasibility study. *International Journal of Developmental Science*, *17*(1/3), 93–110. https://doi.org/10.3233/DEV-230342

Shukla, S., & Otto, C. (2025, February 20). Young voters are powering the rise of Germany's far-right AFD party. *CNN*. https://edition.cnn.com/2025/02/20/europe/afd-germany-elections-young-intl/index.html

Simon, B., Eschert, S., Schaefer, C. D., Reininger, K. M., Zitzmann, S., & Smith, H. J. (2019). Disapproved, but tolerated: The role of respect in outgroup tolerance. *Personality and Social Psychology Bulletin*, *45*(3), 406–415. https://doi.org/10.1177/0146167218787810

Simon, B., & Schaefer, C. D. (2016). Tolerance as a function of disapproval and respect: The case of Muslims. *British Journal of Social Psychology*, *55*(2), 375–383. https://doi.org/10.1111/bjso.12137

Simonsson, O., Goldberg, S. B., Marks, J., Yan, L., & Narayanan, J. (2022). Bridging the (Brexit) divide: Effects of a brief befriending meditation on affective polarization. *PLOS ONE*, *17*(5), Article 0267493. https://doi.org/10.1371/journal.pone.0267493

Smith, J. R., Louis, W. R., Terry, D. J., Greenaway, K. H., Clarke, M. R., & Cheng, X. (2012). Congruent or conflicted? The impact of injunctive and descriptive norms on environmental intentions. *Journal of Environmental Psychology*, *32*(4), 353–361. https://doi.org/10.1016/j.jenvp.2012.06.001

Steinberg, L. (2005). Cognitive and affective development in adolescence. *Trends in Cognitive Sciences*, *9*(2), 69–74.

Steinberg, L., & Monahan, K. C. (2007). Age differences in resistance to peer influence. *Developmental Psychology*, *43*(6), 1531–1543. https://doi.org/10.1037/0012-1649.43.6.1531

Stephan, W. G., & Stephan, C. W. (2016). Intergroup threats. In C. G. Sibley & F. K. Barlow (Eds.), *The Cambridge handbook of the psychology of prejudice* (1st ed., pp. 131–148). Cambridge University Press. https://doi.org/10.1017/9781316161579.007

Svolik, M. W. (2019). Polarization versus democracy. *Journal of Democracy*, *30*(3), 20–32. https://doi.org/10.1353/jod.2019.0039

Tajfel, H., & Turner, J. C. (2004). The social identity theory of intergroup behavior. In J. T. Jost & J. Sidanius (Eds.), *Political psychology* (pp. 276–293). Psychology Press. https://doi.org/10.4324/9780203505984-16

Tang, T., Ghorbani, A., Squazzoni, F., & Chorus, C. G. (2022). Together alone: A group-based polarization measurement. *Quality & Quantity*, *56*(5), 3587–3619. https://doi.org/10.1007/s11135-021-01271-y

Tankard, M. E., & Paluck, E. L. (2016). Norm perception as a vehicle for social change. *Social Issues and Policy Review*, *10*(1), 181–211. https://doi.org/10.1111/sipr.12022

Tausch, N., Birtel, M. D., Górska, P., Bode, S., & Rocha, C. (2024). A post-Brexit intergroup contact intervention reduces affective polarization between Leavers and Remainers short-term. *Communications Psychology*, *2*(1), Article 95. https://doi.org/10.1038/s44271-024-00146-w

Tropp, L. R., O'Brien, T. C., González Gutierrez, R., Valdenegro, D., Migacheva, K., De Tezanos-Pinto, P., Berger, C., & Cayul, O. (2016). How school norms, peer norms, and discrimination predict interethnic experiences among ethnic minority and majority youth. *Child Development*, *87*(5), 1436–1451. https://doi.org/10.1111/cdev.12608

Tyler, M., & Iyengar, S. (2023). Learning to dislike your opponents: Political socialization in the era of polarization. *American Political Science Review*, *117*(1), 347–354. https://doi.org/10.1017/S000305542200048X

Uzun, B., & Lüküslü, D. (2023). Polarisation and youth in Turkey: Young people as gatekeepers or challengers of polarisation? *South European Society and Politics*, *28*(2), 153–175. https://doi.org/10.1080/13608746.2024.2327794

Van De Schoot, R., Lugtig, P., & Hox, J. (2012). A checklist for testing measurement invariance. *European Journal of Developmental Psychology*, *9*(4), 486–492. https://doi.org/10.1080/17405629.2012.686740

Van Deth, J. W., Abendschön, S., & Vollmar, M. (2011). Children and politics: An empirical reassessment of early political socialization. *Political Psychology*, *32*(1), 147–174. https://doi.org/10.1111/j.1467-9221.2010.00798.x

Van Hoorn, J., Crone, E. A., & Van Leijenhorst, L. (2017). Hanging out with the right crowd: Peer influence on risk-taking behavior in adolescence. *Journal of Research on Adolescence*, *27*(1), 189–200. https://doi.org/10.1111/jora.12265

van Noord, J., Turner-Zwinkels, F. M., Kesberg, R., Brandt, M. J., Easterbrook, M. J., Kuppens, T., & Spruyt, B. (2025). The nature and structure of European belief systems: Exploring the varieties of belief systems across 23 European countries. *European Sociological Review*, *41*(1), 143–161. https://doi.org/10.1093/esr/jcae011

Váradi, L., Barna, I., & Németh, R. (2021). Whose norms, whose prejudice? The dynamics of perceived group norms and prejudice in new secondary school classes. *Frontiers in Psychology*, *11*, Article 524547. https://doi.org/10.3389/fpsyg.2020.524547

Velthuis, E., Verkuyten, M., & Smeekes, A. (2021). The different faces of social tolerance: Conceptualizing and measuring respect and coexistence tolerance. *Social Indicators Research*, *158*(3), 1105–1125. https://doi.org/10.1007/s11205-021-02724-5

Verkuyten, M. (2022). *The social psychology of tolerance* (1st ed.). Routledge. https://doi.org/10.4324/9781003335030

Verkuyten, M., Yogeeswaran, K., & Adelman, L. (2019). Intergroup toleration and its implications for culturally diverse societies. *Social Issues and Policy Review*, *13*(1), 5–35. https://doi.org/10.1111/sipr.12051

Verkuyten, M., Yogeeswaran, K., Mepham, K., & Sprong, S. (2020). Interculturalism: A new diversity ideology with interrelated components of dialogue, unity, and identity flexibility. *European Journal of Social Psychology*, *50*(3), 505–519. https://doi.org/10.1002/ejsp.2628

Voelkel, J. G., Stagnaro, M. N., Chu, J. Y., Pink, S. L., Mernyk, J. S., Redekopp, C., Ghezae, I., Cashman, M., Adjodah, D., Allen, L. G., Allis, L. V., Baleria, G., Ballantyne, N., Van Bavel, J. J., Blunden, H., Braley, A., Bryan, C. J., Celniker, J. B., Cikara, M., . . . Willer, R. (2024). Megastudy testing 25 treatments to reduce antidemocratic attitudes and partisan animosity. *Science*, *386*(6719), Article eadh4764. https://doi.org/10.1126/science.adh4764

Wagner, M. (2021). Affective polarization in multiparty systems. *Electoral Studies*, *69*, Article 102199. https://doi.org/10.1016/j.electstud.2020.102199

Wagner, M. (2024). Affective polarization in Europe. *European Political Science Review*, *16*(3), 378–392. https://doi.org/10.1017/S1755773923000383

Wang, C., Hipp, J. R., Butts, C. T., Jose, R., & Lakon, C. M. (2017). Peer influence, peer selection and adolescent alcohol use: A simulation study using a dynamic network model of friendship ties and alcohol use. *Prevention Science*, *18*(4), 382–393. https://doi.org/10.1007/s11121-017-0773-5

Webster, S. W., & Abramowitz, A. I. (2017). The ideological foundations of affective polarization in the U.S. electorate. *American Politics Research*, *45*(4), 621–647.

Whitt, S., Yanus, A. B., McDonald, B., Graeber, J., Setzler, M., Ballingrud, G., & Kifer, M. (2021). Tribalism in America: Behavioral experiments on affective polarization in the trump era. *Journal of Experimental Political Science*, *8*(3), 247–259. https://doi.org/10.1017/XPS.2020.29

You, Z. T., & Lee, S. W. S. (2024). Explanations of and interventions against affective polarization cannot afford to ignore the power of ingroup norm perception. *PNAS Nexus*, *3*(10), Article pgae286. https://doi.org/10.1093/pnasnexus/pgae286

About the Authors

Maor Shani is a post-doctoral researcher at the Department of Developmental Psychology at Osnabrück University and a research fellow at the Faculty of Social Sciences at Ariel University. His work focuses on intergroup conflicts, interventions to improve intergroup relations, antisemitism, and coping with chronic illness.

Marjorie Berns is a doctoral candidate in the Department of Social Psychology at the Osnabrück University, conducting research on racism in sports.

Lucy Bergen earned a bachelor's degree in psychology at Osnabrück University and her research focuses on interpersonal relationships, social dynamics, and opinion formation during adolescence.

Stefanie Richters is a PhD student at the Department of Developmental Psychology at the University of Osnabrück in Germany. Her research interests include network dynamics in schools, (cyber-)bullying, and social influence through norms.

Kristina Krämer (she/her) is a research assistant at Osnabrück University's Department of Developmental Psychology, with research interests in interventions targeting prosocial behavior among adolescents. She also qualifies as a child and adolescent psychotherapist.

Sophie de Lede (she/her) is a predoctoral researcher at Osnabrück University's Department of Developmental Psychology, with research interests in interpersonal and intergroup relations, conspiracy beliefs, and climate action.

Maarten van Zalk is a full professor of developmental psychology at Osnabrück University who focuses on adolescent development of intergroup relations, social relationships, and antisocial behavior.

Social Inclusion is a peer-reviewed open access journal which provides academics and policymakers with a forum to discuss and promote a more socially inclusive society.

The journal encourages researchers to publish their results on topics concerning social and cultural cohesiveness, marginalized social groups, social stratification, minority-majority interaction, cultural diversity, national identity, and core-periphery relations, while making significant contributions to the understanding and enhancement of social inclusion worldwide.

cogitatio

www.cogitatiopress.com/socialinclusion