Journal of Personality and Social Psychology:
Interpersonal Relations and Group Processes

ISSN: 0022-3514

2024, Vol. 126, No. 4, 624–642
https://doi.org/10.1037/pspi0000454

How Perceived Polarization Predicts Attitude Moralization (and Vice Versa): A Four-Wave Longitudinal Study During the 2020 U.S. Election

Chantal D'Amore, Martijn van Zomeren, and Namkje Koudenburg
Department of Social Psychology, University of Groningen

Within structurally polarized and dynamic contexts, such as the U.S. 2020 presidential elections, the moralization of individuals' attitudes on a specific topic (e.g., climate policy) can dangerously escalate disagreements between groups into zero-sum conflict. However, limited knowledge exists regarding the factors that influence individuals' tendency to moralize their attitudes over time, and what the role of structural polarization is in this psychological process. Our objective is to test a theoretically integrative model of when and how perceived polarization is related to attitude moralization over time within the polarized context of the U.S. 2020 presidential elections and explore reciprocal feedback loops to understand the dynamic relationship between polarization and moralization over time. Our model predicts that, when repeatedly faced with outgroup expressions in the news, individuals' perceptions of polarization will predict within-person attitude moralization over time via strengthening their value-protective responses to these expressions (i.e., perceiving dyadic harm and experiencing negative moral emotions toward the outgroup). To test our model, we conducted a four-wave, 4-month longitudinal study among Biden supporters (N = 1,236) and Trump supporters (N = 617). The results of the within-person analyses generally supported the model's hypotheses across both samples and various attitude topics. Furthermore, cross-lagged structural equation models explored reciprocal influences, revealing positive feedback loops between structural polarization and attitude moralization over time. Our findings thus indicate that perceived polarization strengthens attitude moralization (and vice versa) over time—a dynamic process that helps to explain how nonmoralized conflict between groups can evolve into zero-sum conflict during periods of intense polarization.

Keywords: attitude moralization, perceived polarization, network homogeneity, perceived dyadic harm, negative moral emotions

Supplemental materials: https://doi.org/10.1037/pspi0000454.supp

> What was once a positive-sum negotiation becomes a zero-sum war.
> —(Klein, 2020, p.13)

The 2020 U.S. presidential election campaign was characterized by strong structural polarization, marked by clear "us" versus "them" disagreement (Koudenburg & Kashima, 2022), and outright moral conflict, exemplified by events like the 2021 Capital Hill riot. Throughout the campaign, Trump supporters and Biden supporters fervently disagreed on specific topics, such as COVID-19, climate change, and the new Supreme Court Justice. More significantly, they clashed on what they believed to be morally right and wrong (e.g., opposing candidates accusing the other of immoral intent and calling them names such as "enemy of the state" or "climate arsonist"). This shift from mere attitudinal disagreement to moral conflict reflects a potentially dangerous dynamic for democratic societies, because the stronger people moralize attitudes on specific topics, the more motivated they become to actively protect their convictions and underlying values (Ellemers et al., 2019; Skitka et al., 2021). Unlike attitudes more generally, such moralized attitudes (i.e., *moral convictions*) are experienced as absolute facts about the world that, when threatened or doubted by others, require protection at almost all costs and by any means necessary (e.g., vigilantism; Skitka, 2002; for reviews, see Kovacheff et al., 2018; Skitka et al., 2021; Skitka & Morgan, 2014). To illustrate this potential danger, moral convictions have been linked with opposition against constructive negotiation or compromise, intolerance, hostility, and willingness to reject the rule of law or accept the use of violence against opponents (e.g., Cole Wright et al., 2008; Mooijman et al., 2018; Ryan, 2017; Simonsen & Bonikowski, 2022; Zaal et al., 2011, 2017). As Klein puts it in the opening quote, attitude moralization thus facilitates a concerning shift from positive-sum negotiations toward a zero-sum war.

This article was published Online First February 22, 2024.
Chantal D'Amore https://orcid.org/0000-0002-7151-3888
The authors have no conflicts of interest to disclose. Data, materials, R-Scripts, and additional online materials are openly available at the project's Open Science Framework page (https://osf.io/f9p4e/; D'Amore, 2024).

Chantal D'Amore played a lead role in data curation, formal analysis, investigation, methodology, project administration, software, validation, visualization, writing–original draft, and writing–review and editing and an equal role in conceptualization. Martijn van Zomeren played a lead role in supervision, a supporting role in investigation, methodology, writing–original draft, and writing–review and editing, and an equal role in conceptualization and validation. Namkje Koudenburg played a supporting role in conceptualization, formal analysis, investigation, methodology, validation, visualization, writing–original draft, and writing–review and editing and an equal role in supervision.

Correspondence concerning this article should be addressed to Chantal D'Amore, Department of Social Psychology, University of Groningen, Grote Kruisstraat 2/1, 9712 TS Groningen, The Netherlands. Email: c.damore@rug.nl

Surprisingly little is known, however, about what makes people develop or strengthen moralized attitudes within structurally polarized contexts. Attitude moralization is defined as the psychological process through which an individual subjectively connects their attitude on a specific issue with their fundamental beliefs about right and wrong (i.e., developing or strengthening a moral conviction; Skitka et al., 2021). Reflecting this subjective nature, research shows that individuals differ substantially in the degree to which they have moralized their attitude on political topics (e.g., Rhee et al., 2019; Ryan, 2014; Skitka et al., 2021)—even when it comes to political topics that are considered prototypically "moral" on theoretical grounds (e.g., abortion, capital punishment). Concerning antecedents, the prevailing view in the literature assumes that attitude moralization results from topic-specific triggers (e.g., "moral shocks" relevant to abortion or meat-eating; Feinberg et al., 2019; Wisneski & Skitka, 2017), which does not take into account the potential of key aspects of polarized contexts to feed into the moralization process. More recently, however, experimental research identified group-based triggers for attitude moralization emerging within polarized contexts, namely political outgroup expressions that are perceived as involving strong intentional harm against the ingroup (i.e., strong *dyadic harm*; D'Amore et al., 2022; Schein & Gray, 2018). We, therefore, build on these insights to develop and test an integrative social psychological model of when and how individuals come to strengthen their moral convictions over time within the dynamic and structurally polarized context of the U.S. 2020 election campaign.

The current research breaks new ground through achieving three specific aims. First, we seek to move beyond previous work by focusing on the *within-person* drivers of attitude moralization within a dynamic polarized context over time. Specifically, rather than examining whether exposure to specific stimuli (e.g., outgroup dyadic harm) functions as a moralization trigger, this study aims to test whether subjective increases in individuals' own responses to outgroup expressions over time drive increases in their own moral convictions over time. This first aim is pivotal because although attitude moralization is by definition a within-person process that unfolds over time, research has seldomly analyzed its potential predictors at this within-person level (e.g., Brandt et al., 2015; Rhee et al., 2019; Wisneski et al., 2020). Second, we seek to examine whether, given repeated exposure to outgroup expressions, individuals' perceptions of polarization feed into this moralization process about and across specific topics. This is key because it moves us beyond the narrow focus in the existing literature on topic-specific predictors of moralization by integrating contextual factors. Third, zooming out from the within-person moralization process, we explore the dynamic reciprocal relationship between perceived polarization and attitude moralization longitudinally. Considering that little is still known about this, we take advantage of the collected longitudinal data to explore whether the model can be expanded to include feedback loops. We specifically explore whether both attitude moralization and perceived polarization reinforce each other over time (i.e., positive feedback loops). This question is of central relevance to understanding the bigger picture of how attitudinal disagreement transforms into moral conflict in society and may therefore offer valuable directions for future research.

Our model conceptually integrates theories of polarization, dyadic harm, value-protection, and moral conviction (Koudenburg & Kashima, 2022; Schein & Gray, 2018; Skitka et al., 2021; Tetlock et al., 2000; van Zomeren et al., 2023). In line with D'Amore et al. (2022) we expect that strong *perceptions of dyadic harm* (i.e., perceived intentional actor causing damage to a vulnerable victim; Schein & Gray, 2018) and strong *negative moral emotions* (e.g., anger, disgust; Skitka et al., 2021) felt toward outgroup expressions trigger attitude moralization within polarized contexts. Extending this to within-person drivers of this process, increased attitude moralization occurs when an individual comes to perceive stronger dyadic harm in, and experience stronger negative moral emotions toward, outgroup expressions than before. Connecting those drivers to the structural context, increasing perceptions of polarization (i.e., perceptions of polarization in society, and as reflected in the political homogeneity of one's network) over time predict increases in both drivers of moralization over time and in this way facilitate increased attitude moralization. This is because such perceptions facilitate that individuals use a group-based lens through which to interpret and respond to politically relevant news (Koudenburg & Kashima, 2022). Thus, we expect that when people perceive stronger polarization than they did before, this will lead them to perceive stronger dyadic harm in, and respond with stronger negative moral emotions to outgroup expressions than they did before, and hence their attitude moralization on specific topics strengthens.

To test our model and explore feedback loops, we designed and conducted a four-wave longitudinal study spanning over 4 months of the U.S. 2020 election campaign. We measured both Biden and Trump supporters' perceptions of structural polarization (both in society and as reflected in one's network) at four different time points, in order to connect the within-person changes in these perceptions to within-person changes in their perceptions of dyadic harm, negative emotions, and attitude moralization on specific topics from one time point to another. We achieved this by actively reminding participants in each wave of recent potentially harmful outgroup expressions covered by the news, which enabled us to measure individuals' wave-specific perceptions of dyadic harm and experience of negative moral emotions with explicit reference to those outgroup expressions, and thus to capture within-person changes in both drivers of moralization over time. Our research design also allowed us to test our model on two different samples (i.e., Trump and Biden supporters) in order to assess whether the hypothesized processes are generalizable across the political divide. Additionally, for each sample, we selected three specific topics relevant to potential attitude moralization in the election context, aiming to assess the generalizability of findings across various salient topics. Taken together, these longitudinal, within-person, multisample, and multitopic aspects of our research design uniquely enabled achieving our aims and testing our model's hypotheses about within-person drivers and predictors for attitude moralization against the backdrop of the U.S. 2020 election, while also offering scope for exploring potential feedback loops.

An Integrative Social Psychological Model of Attitude Moralization Within Polarized Contexts

In this first section, we outline our conceptual model to test (1) a within-person process and drivers of attitude moralization in polarized contexts, (2) the role of perceived polarization as a predictor, and (3) the exploration of potential feedback loops between attitude moralization and perceptions of polarization.

Attitude Moralization Within Persons and Its Psychological Drivers

Attitude moralization is defined as the psychological process by which a given attitude becomes increasingly connected to one's core convictions and beliefs about fundamental right and wrong (e.g., Skitka et al., 2021) and marks a profound psychological shift within-individuals. An individual's anti- or proabortion attitude, for example, may initially reflect an opinion on this specific topic. However, through the moralization process, this attitude evolves to reflect fundamental beliefs about right and wrong (i.e., it becomes grounded in core values such as "prochoice"; see also *moral piggybacking*; e.g., Feinberg et al., 2019; Rozin, 1999). Attitude moralization can therefore be conceptualized as part of a *value-protective* process (Tetlock et al., 2000) in response to outgroup expressions perceived as intentionally harmful (D'Amore et al., 2022), serving to protect those values from perceived and anticipated future outgroup attacks. Indeed, perceiving outgroup expressions as involving strong intentions to harm innocent victims subjectively transforms the outgroup into a dangerous enemy that aims to destroy core values in society (e.g., equality/autonomy; D'Amore et al., 2022). This triggers a highly emotional value-protective response that involves strong negative moral emotions such as anger and disgust, which push the topic-specific attitude into the moral domain (Brandt et al., 2015; D'Amore et al., 2022; Leal et al., 2023; Wisneski & Skitka, 2017). Figure 1 visualizes how our model proposes that individual increases in attitude moralization of a specific topic are likely driven by individual increases in their value-protective responses to outgroup expressions over time. Specifically, our integrative model identifies two important psychological drivers for attitude moralization that reflect those value-protective responses: perceptions of dyadic harm in outgroup expressions, and the experiencing of negative moral emotions.

Perceived Dyadic Harm. The first psychological driver features perceptions of outgroup *dyadic harm*. Dyadic harm is defined as the subjective perception of an intentional actor causing damage to a vulnerable victim (Schein & Gray, 2018). This specific type of perceived harm is thought to be uniquely linked to the moral domain and hence sets it apart from traditional or general perspectives on harm (e.g., Haidt, 2012). The theory of dyadic morality (Schein & Gray, 2018; see also Gray et al., 2022) places two causally connected minds at the center of moral judgment (the dyadic template), suggesting that moralization emerges when topic-relevant expressions are subjectively perceived as involving a combination of (a) the *suffering* of victims, (b) the *conscious intention* of an actor, and (c) their *dyadic causality*. For example, in a series of experiments in the context of a heated political debate in the Netherlands, when an outgroup announced their willingness to use intimidation, provocation, and violence against the ingroup (rather than understanding, tolerance, and solidarity) to change a Dutch tradition that includes blackface, this led supporters of this tradition to perceive stronger dyadic harm and, in turn, strengthened the moralization of their attitude on this topic (D'Amore et al., 2022). The dyadic template thus conceptually helps to understand why subjective perceptions of a political outgroup intentionally inflicting harm on a political ingroup can be a psychological driver of moralization about specific topics within polarized contexts (D'Amore et al., 2022).

Negative Moral Emotions. This second driver has also received empirical support. For example, the same studies by D'Amore et al. (2022) found that moralization occurred because negative moral emotions such as anger were experienced in response to the outgroup's dyadic harm expression. A different line of experimental studies found that attitude moralization (e.g., against sexism) increased when expressions from an immoral outgroup were perceived to violate ingroup values, again mediated by strong emotional responses (Leal et al., 2023). Such findings fit with the idea that the strength of emotions felt in relation to the topic of moralization explains variance in its occurrence (Brandt et al., 2015; Feinberg et al., 2019; Wisneski & Skitka, 2017; for a review, see Skitka et al., 2021).

In sum, our model proposes two important within-person drivers for attitude moralization: perceptions of dyadic harm and the experience of negative moral emotions. The model predicts that, when repeatedly exposed to recent outgroup expressions in the news within polarized contexts, within-person increases in perceived dyadic harm (H1.1) and in negative moral emotion (H1.2) over time predict within-person increases in attitude moralization. However, this reasoning does not yet take into account the potentially predictive influence of structural polarization, which is discussed below.

How Structural Polarization Predicts Drivers of Moralization Within Individuals

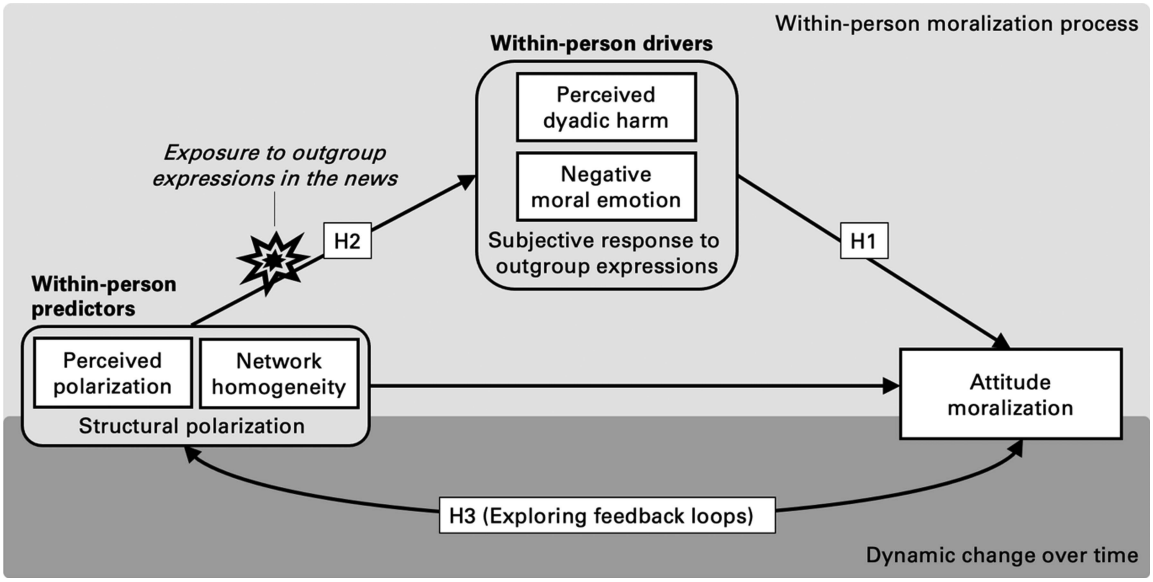
Figure 1 visualizes how increased structural polarization feeds into the process of moralization within-individuals via predicting increases in individuals' negative moral emotions and perceptions of dyadic harm when considering outgroup expressions. We define structural polarization as the structural divisions existing between groups in society. Perceptions of polarization thus refer to the extent to which people believe that different groups in society structurally stand in direct opposition with respect to their positions on relevant political topics (Koudenburg & Kashima, 2022). Moreover, structural polarization involves the interplay of intergroup and intragroup processes at different levels. That is, while political differences between groups at the societal level become larger, political differences between individuals at the social network level become smaller (e.g., political segregation; Goldenberg et al., 2020; and political sorting; Brown & Enos, 2021). We propose that structural polarization predicts the strengthening of moralization drivers over time in both ways: through individuals' increasing perceptions of polarization in society[1] (Figure 1) and through individuals' embeddedness in an increasingly politically homogeneous network (Figure 1)—as outlined below.

Perceived Polarization. Perceiving increasing polarization in society implies perceiving growing structural differences and attitudinal conflict between (partisan) groups (e.g., Democrats and Republicans). Based on the theory and research in the social identity tradition (for a review, see Hornsey, 2008), this perception leads individuals to subjectively categorize themselves as members of

[1] While previous research has indicated that perceptions of polarization do not always align with existing divisions in society (Lelkes, 2016; Levendusky & Malhotra, 2016a), such perceptions tend to have stronger psychological consequences (Enders & Armaly, 2019; Koudenburg & Kashima, 2022).

Figure 1

Conceptual Model Connecting Structural Polarization With Attitude Moralization



Note. This figure visualizes the hypothesized within-person moralization process over time when repeatedly exposed to outgroup expressions in the news (light-grey box, top) and the explored feedback loops of dynamic change between attitude moralization and perceived polarization (dark-grey box, bottom). When repeatedly exposed to outgroup expressions in the news, within-person increases in perceptions of structural polarization over time will serve as predictors of attitude moralization (H2) via predicting within-person increases in responses to the outgroup expressions over time (H1). H1 = Hypothesis 1; H2 = Hypothesis 2; H3 = Hypothesis 3.

their political ingroup in situations where politics are salient (e.g., Brady et al., 2020; Van Bavel & Packer, 2021). Consequently, stronger perceptions of polarization make it more likely for individuals to evaluate the relevant outgroup and their expressions through a negative lens (Tajfel & Turner, 1986) and to feel motivated to defend their ingroup's moral image and core values from outgroup attacks (Brady et al., 2020; Ellemers et al., 2019; Ellemers & van den Bos, 2012). Consistent with this, perceiving stronger polarization between groups has been linked to more negative affective evaluations of political opponents (Enders & Armaly, 2019), and a stronger sense of relational threat and negative emotions toward opponents (Koudenburg & Kashima, 2022). Moreover, the more structural differences people perceive to exist between a given (political) outgroup and ingroup, the more motivated they tend to become to make the ingroup appear (morally) superior by comparison, for example, by interpreting the outgroup's intentions in a negative moral light (Brady et al., 2020). This is in line with the idea that stronger perceptions of structural polarization in society strengthen both psychological drivers of moralization within-individuals: namely, perceived dyadic harm and negative moral emotions experienced in response to the political outgroup and their expressions in the news.

Perceived Network Homogeneity. Our conceptual model further posits that individuals' embeddedness in an increasingly politically *homogeneous* network, as opposed to a heterogeneous one, can predict increases in both moralization drivers in a comparable manner. This is because being surrounded by, and interacting with, like-minded individuals likely reinforces negative evaluations and value-protective responses toward the relevant outgroup. For example, research suggests that individuals within more politically homogeneous social network are more likely to engage in motivated reasoning and maintain a negative bias toward opponents (Levitan & Visser, 2008, 2009; Levitan & Wronski, 2014). Likewise, studies have demonstrated that reduced political disagreement within one's network was associated with heightened affective polarization, that is, stronger negative affective evaluations of the outgroup-candidate compared to the ingroup-candidate (Lupton et al., 2015). Furthermore, perceived moral homogeneity (i.e., homogeneity in moral concerns) within networks has been associated with increased intentions to engage in radical (illegal or violent) political action against opponents (Atari et al., 2022; see also Mooijman et al., 2018). This fits with the idea that more homogeneous networks strengthen negative emotions and perceptions of dyadic harm in response to relevant outgroups, and thus predict within-person drivers of attitude moralization within polarized contexts.[2]

In summary, our model predicts that when individuals perceive more polarization between groups or are embedded in more homogeneous networks than before, they are likely to perceive stronger dyadic harm and experience stronger negative moral

[2] Beyond the focus of the present article, potential mechanisms explaining these effects of network homogeneity highlighted elsewhere include selective exposure to negative outgroup information and moral-emotional content (e.g., Brady et al., 2020) and shared reality regulation and social validation through interpersonal interaction within homogeneous groups (e.g., Echterhoff & Higgins, 2017; Koudenburg, 2018; Price et al., 2006).

emotions in the face of potentially harmful outgroup expressions than before. When repeatedly exposed to outgroup expressions about specific topics in the news (e.g., COVID mask-wearing), heightened perceptions of polarization and network homogeneity are expected to predict the strengthening of individuals' value-protective responses to these outgroup expressions, which, in turn, are expected to drive increased moralization of attitudes on the relevant topic within-individuals. Thus, our model's specific hypotheses are that within-person increases in perceived polarization (H2.1) and network homogeneity (H2.2) predict within-person increases in perceived dyadic harm and negative moral emotion to the outgroup, and thereby (i.e., indirectly) predict within-person increases in attitude moralization.

Exploring Dynamic Feedback Loops Between Polarization and Moralization Over Time

Figure 1 (bottom part) visualizes a dynamic feedback loop between structural polarization and attitude moralization over time. Moving beyond the within-person moralization processes hypothesized above, this loop suggests that the strength of a moral conviction at a given time point predicts increases in perceived polarization at the subsequent time point (and vice versa). If the data provide initial support for a positive feedback loop, this could offer a valuable direction for future research to understand how exactly the two processes interact—and how this understanding may be used to predict and potentially prevent the escalation of conflict between groups in society (van Zomeren et al., 2023).

The presence of positive feedback loops between moralization and polarization is plausible for at least two reasons. First, previous research suggests that the more individuals moralize their attitude on a specific topic, the more they tend to perceive people who share their attitude on this specific topic as fundamentally right and those who disagree as fundamentally wrong (Skitka et al., 2021). As individuals tend to perceive their moral convictions as objective truths, strong moral convictions tend to make people perceive debates around this specific topic in terms of "us-versus-them," "right-versus-wrong," with little room for nuance. Second, the more people moralize, the more they begin to comprehend differences between groups not only in terms of (topic-)specific attitudes but also in terms of more abstract moral values and beliefs. Indeed, because individuals' sense of morality significantly influences their evaluations of other people and groups (e.g., Ellemers et al., 2019), stronger moral convictions about specific topics likely fuel stronger perceptions of the political ingroup and outgroup as fundamentally different and immoral. We therefore explore whether stronger moral convictions at a given time point predict stronger increases in perceived polarization at a subsequent time point (Hypothesis 3), suggesting a dynamic reciprocal relationship between these two processes.

Method

We empirically tested our model through a four-wave longitudinal study set in the dynamic and structurally polarized context of the 2020 U.S. presidential elections. Our approach advances prior work in at least three significant ways. First, while previous experimental research manipulated either of both triggers between persons (i.e., negative moral emotions or perceived dyadic harm), it provided limited insights into their potential as within-person drivers of moralization over time. Therefore, it remains unclear whether the strengthening of individuals' own psychological (value-protective) responses to outgroup expressions predicts the extent to which people will moralize their attitudes about specific topics upon repeated exposure. The present longitudinal study enables answering this question. Second, if these within-person increases in perceived dyadic harm and negative moral emotions indeed drive further moralization, then the question arises: what prompts individuals to strengthen their own responses to outgroup expressions in the first place? We test whether and how individuals' perceptions of polarization predict the strengthening of both moralization drivers within-individuals.[3] Third, this research uniquely explores potential reciprocal influences between perceived polarization and attitude moralization over time.

Situated in the context of the U.S. 2020 elections, we designed and conducted a four-wave longitudinal study to test our model's predictions about and across different key topics (e.g., Mask-Wearing, Paris Climate Agreement, Local Law Enforcement, Supreme Court Justice), among and across different samples (i.e., Biden and Trump supporters). This election context was highly suited to test our model because it involved (a) structural polarization between two groups across various specific topics, (b) natural variation in perceptions of such polarization both between individuals and within-individuals over time, and (c) a continuous flow of naturally emerging news messages covering recent expressions by groups on both sides across a range of specific topics (e.g., consider an act of candidate Trump mocking Joe Biden for wearing masks or that of candidate Biden calling Trump a "climate arsonist"). These news messages featured potentially harmful outgroup expressions about different topics (i.e., mask-wearing and climate policy, respectively) and involved the same outgroup representing the broader polarized context at hand. All of this allowed us to use naturally emerging information about the relevant outgroup as real-time stimuli in our survey to assess individuals' responses in terms of dyadic harm and emotions, and to connect those responses with individuals' perceptions of polarization on the one hand and attitude moralization on the other. In terms of our statistical analysis plan, we used multilevel regression analyses to test our two hypotheses about purely within-person processes (Hypothesis 1 and Hypothesis 2), and we used cross-lagged structural equation analyses to explore dynamic change and reciprocal influences between polarization and moralization over time (Hypothesis 3).

Transparency and Openness

Below, we describe how we determined our sample size (power analysis), outline voluntary dropout as well as exclusion criteria and rates, and evaluate the final sample sizes for each topic-specific (sub)sample in comparison with the estimated required sample size. All materials, anonymized data, and R-Scripts used for this article are available on Open Science Framework (link: https://osf.io/f9p4e/). This study was not preregistered online.

[3] If anything, D'Amore et al. (2022) showed that people who perceived stronger polarization than others were more likely to have stronger moral convictions than others (i.e., a between-subject association).

Sample Size Determination

To estimate the minimum required sample size, we conducted a priori power analysis in G*Power (Version 3.1.9.3) and used the smallest effect size ($f = 0.071$) found in a study on attitude moralization using similar measures as in the present study (D'Amore et al., 2022) as a conservative estimate for within-individual change.[4] To obtain a power of at least 80%, given an estimated correlation between repeated measures of intraclass correlation coefficient, $ICC = .71$ (D'Amore et al., 2022) and $\alpha = .05$, the computed sample size for a repeated-measures (within-subject) test with four repeated measures was $N = 159$ for each topic-specific (sub)sample. Taking into account participant dropout in longitudinal research (e.g., near 60% in MTurk samples; Berinsky et al., 2012; Brandt et al., 2015), and participant exclusion by a number of criteria (e.g., ensuring high-quality data and repeated observations for each participant; see Participants and Procedure below), we planned to collect a substantially larger baseline number (at Wave 1) of 500 participants at in each sample (i.e., $N = 500$ Biden supporters in Sample 1 and $N = 500$ Trump supporters in Sample 2), who were then invited to participate in each of the subsequent waves.

Participants and Procedure

Participants were all U.S. residents who reported at least some support for Joe Biden (Sample 1) or for Donald Trump (Sample 2) within the 2020 presidential election in the United States. This means that only the participants who did not indicate a preference for one of both presidential candidates were excluded after the first survey question. They were recruited via an online provider (MTurk) and preselected based on extensive screening to ensure good data quality (i.e., rigorous checks for bots and nonattentive participants). Participants received monetary rewards for their participation in each wave ranging between \$0.50 at Wave 1 to \$1.50 at Wave 4 (with Wave 2: \$0.90, and Wave 3: \$1.50). In total, the survey took approximately 20 min to complete at Wave 1, and 14 min at each subsequent measurement wave. This research project was approved by the Ethics Committee of the Faculty of Behavioral and Social Sciences at the University of Groningen.

Supplemental Materials Figure A1 (p. 3) shows a complete flowchart of participants, dropout rates, and exclusion by wave. At Wave 1 (out of four), the combined samples initially included 1,075 observations, which was reduced to a combined sample of 931 participants after exclusion of (a) duplicate participant response (i.e., participants who completed the same survey twice at T1—an error which was later resolved; $n = 65$), and (b) nonsensical responses to the open questions ($n = 49$). Consequently, the total samples at Wave 1 included 485 Biden supporters (Sample 1) and 446 Trump supporters (Sample 2). These participants were all invited to participate in each of the following three waves (response rates relative to T1: T2 = 56.9%, T3 = 54.8%, T4 = 49.2%) and were included for analyses when they participated in at least two out of the four waves and consistently indicated support for either Biden or Trump across those waves.[5] This resulted in a final number of 355 Biden supporters in Sample 1, 41% male, 57% female (1% other); Mage = 45.31, $SD = 13.58$; education level: low (3%), intermediate (72%), high (25%), and 186 Trump supporters in Sample 2, 44% male, 56% female; Mage = 46.73, $SD = 13.80$; education level: low (9%), intermediate (71%), high (17%).

Data Collection

Data collection for both samples involved a total of four waves, each measured at a meaningful time point within the lead-up to, and aftermath of, the presidential election. The baseline measurement (T1) took place 1 month before the election (September 26–October 2, 2020), with already salient polarization (see Tables 1 and 2, for variable descriptives by wave, for Samples 1 and 2, respectively). The second wave (T2) was collected right before election day (October 26–November 2), the third wave (T3) was 1 month after the election (November 24–30), and the final measurement (T4) was right after the inauguration day and the Capitol Hill riot (January 21–28).

Selected Political Topics

To test our model's predictions and assess its replicability across multiple political topics, we selected four specific topics that were widely discussed in the lead-up to the election (e.g., in the media). Moreover, in line with our specific design, each of the selected topics was required to be formulated as a concrete political topic (rather than abstract ideas such as "Build Back Better"), and to align with the currently existing political divide between Biden supporters and Trump supporters as indicated by the expected attitude positions of the relevant political groups. Based on these criteria, we included the following topics across all four waves: (a) *Supreme Court Justice* (whether or not the confirmation of a new justice should take place before election day), (b) *Mask-Wearing* (whether or not the United States should implement a mandate for public places), (c) *Climate* (whether or not the United States should rejoin the Paris Agreement; Biden sample), and (d) *Local Law Enforcement* (Trump sample). In terms of each sample's expected attitude positions, Biden supporters were expected to report disagreement for the first topic (Justice) and agreement for the second and the third topics (Mask, Climate), whereas the opposite was expected for Trump supporters for the first two topics (Justice, Mask) with expected agreement on the third topic (Law). As such, this study included three topics in each sample across the four waves, which allowed us to test our model across six cases and thus assess its replicability across different topics and samples.

[4] The validity of this approach is limited by the fact that this reference study (D'Amore et al., 2022) differs from our current examination in important ways (e.g., design). In the absence of better suited alternatives in the existing literature (with the required information on effect size and *ICC* available), we also computed the required sample size in an alternative way—using a short-cut for a small-sized effect ($f = 0.10$) and the default correlation ($ICC = .50$)—which suggested a comparable result of $N = 138$ (i.e., slightly smaller than $N = 159$). However, to account for the slight reduction in actual power resulting from our nonbalanced design (e.g., average number of repeated measures = 3.50 in the Biden sample), we reasoned that the use of a slightly larger required sample size (i.e., $N = 159$) is appropriate here.

[5] The first exclusion criterion (i.e., single-time participation) lead to the exclusion of $N = 107$ participants in the Biden Sample and $N = 214$ in the Trump Sample. The second exclusion criterion (i.e., inconsistent Biden/Trump support across time) lead to the exclusion of $N = 23$ participants in the Biden Sample and $N = 46$ in the Trump Sample.

Table 1
Variable Descriptives (Biden Sample): Intraclass Correlation Coefficients and Means (SD) by Time

Variable	*ICC*	Time 1	Time 2	Time 3	Time 4
Perceived polarization	.640	4.21 (0.64)	4.19 (0.68)	4.13 (0.69)	4.18 (0.66)
Network homogeneity	.726	3.93 (0.73)	3.90 (0.71)	3.99 (0.71)	4.02 (0.73)
Moral conviction					
Climate	.654	3.75 (1.16)	3.71 (1.17)	3.71 (1.23)	3.81 (1.14)
Justice	.582	3.76 (1.20)	3.71 (1.17)	3.52 (1.24)	3.43 (1.26)
Masks	.609	3.95 (1.11)	4.12 (1.00)	4.11 (1.03)	4.11 (1.06)
Perceived dyadic harm					
Climate	.645	3.64 (1.16)	3.50 (1.18)	3.56 (1.25)	3.37 (1.22)
Justice	.651	3.88 (1.21)	4.02 (1.09)	3.86 (1.15)	3.54 (1.29)
Masks	.665	3.92 (1.07)	3.94 (0.98)	3.90 (1.05)	3.88 (1.09)
Negative moral emotion					
Climate	.686	3.82 (1.17)	3.74 (1.20)	3.65 (1.26)	3.69 (1.24)
Justice	.700	4.00 (1.15)	4.06 (1.09)	3.76 (1.20)	3.68 (1.31)
Masks	.681	4.20 (1.01)	4.26 (0.98)	4.18 (1.00)	4.23 (1.05)

Note. Sample sizes vary by time point (N_{T1} = 355; N_{T2} = 303; N_{T3} = 302; N_{T4} = 275) and topic: Climate (N_{T1} = 345; N_{T2} = 296; N_{T3} = 293; N_{T4} = 266), Justice (N_{T1} = 313; N_{T2} = 269; N_{T3} = 266; N_{T4} = 242), Masks (N_{T1} = 347; N_{T2} = 296; N_{T3} = 294; N_{T4} = 268). *ICC* = intraclass correlation coefficient; T = Time.

Topic-Based Sample Sizes

All participants completed the total survey including all three topics. However, topic-based sample sizes vary as they depend on the number of participants who aligned with the position of their political group within our sample. Indeed, a key assumption of our model is that its predictions apply to an intergroup context in which the opinion of the outgroup (featured in the news messages) opposes, or at least does not support, respondents' opinion on this specific topic.[6] Thus, given that both the topic-based cues to dyadic harm from outgroup expressions in the news, as well as the operationalizations of the topic-specific variables (i.e., perceived dyadic harm, negative moral emotion, and attitude moralization), are tailored to the direction of participants' topic-specific attitude, this study excluded those who had the opposite attitude on a specific topic[7] (see D'Amore et al., 2022, for a similar approach).

Biden Supporters (Sample 1). In the Biden sample, most participants had the assumed attitudes on the three specific topics: *Climate* (97%, $N = 345$), *Justice* (88%, $N = 313$), and *Masks*: (98%, $N = 347$). Hence, in line with our statistical power analysis, the sample sizes for each topic in the Biden sample are sufficient.

Trump Supporters (Sample 2). In the Trump sample, most participants had the assumed attitudes on two out of the three specific topics: *Law* (98%, $N = 182$), and *Justice* (96%, $N = 178$). In contrast to our expectation, however, only a minority of participants in the Trump sample had the assumed attitude on the *Masks* topic (38%, $N = 70$). Thus, with regard to statistical power, the final sample sizes for the topics *Law* and *Justice* in the Trump sample are sufficient (i.e., larger than the required $N = 159$) while the sample for the *Mask* topic is clearly under-powered.[8] Hence, in the Trump sample, the results for the *Mask* topic should be interpreted with caution: Below, the results will be reported without further interpretation.

Procedure

In each wave, the procedure and order of measurements[9] were as follows. Participants first indicated whether and how strongly they supported the two presidential candidates in the 2020 election (5-point scale, from 1 = *strong Trump supporter*, to 3 = *neither Trump nor Biden supporter*, to 5 = *strong Biden supporter*). Next, they completed a set of measurements tapping into their perceptions of polarization and network political homogeneity. Following this, participants were presented with a set of topic-specific measurements: They first indicated their attitude on the specific topic, then were presented with topic-specific outgroup expressions in the news (see next section) and indicated their responses in terms of perceived dyadic harm and negative moral emotions, and then reported on the moralization of their topic-specific attitude. Subsequently, participants completed this same set of measurements another two times for the two remaining political topics (presented in a randomized order). Finally, they were asked to indicate some demographic information.

Potentially Harmful Outgroup Expressions in the News. In each wave, participants were presented with topic-specific news messages that emerged in between two consecutive waves, carefully selected to represent potentially harmful expressions

[6] There are two main reasons for this. First, our theoretical rationale implies that perceptions of dyadic harm and negative moral emotions should predict moralization when they are explicitly tailored to a meaningful outgroup in relation to a specific topic, as featured in the outgroup expressions in our research. This is not the case if participants share the opinion of their political outgroup rather than their ingroup on a specific topic (e.g., Trump supporters who support a mask-wearing mandate), and hence the outgroup expressions and measurements they received for those variables were not meaningful. Second, given that attitude moralization is conceptualized as individual change in moral conviction around one's specific attitude position on a topic, individual change in moral conviction only represents attitude moralization if the attitude that is moralized remained similar in meaning.

[7] This means that participants were only excluded if they had the opposite attitude (i.e., scoring a "1 = *strongly oppose*" or a "2 = *oppose*" on a 5-point scale), and included if they had either a supportive or neutral attitude (i.e., scoring either a "3 = *neither oppose nor support*", or a "4 = *support*", or a "5 = *strongly support*").

[8] A post hoc analysis suggested that the achieved power to detect a small-sized effect (of $f = 0.10$) was less than 50% (i.e., 48%).

[9] Besides the measurements reported below, this research included some (wave-specific) measurements for exploratory purposes that were not used for the present research (for an overview, see Supplemental Materials, Table D1, pp. 21–24).

Table 2
Variable Descriptives (Trump Sample): Intraclass Correlation Coefficients and Means (SD) by Time

Variable	*ICC*	Time 1	Time 2	Time 3	Time 4
Perceived polarization	.612	4.04 (0.67)	4.08 (0.66)	4.13 (0.69)	3.61 (0.93)
Network homogeneity	.741	3.91 (1.01)	3.91 (0.71)	3.91 (0.72)	3.89 (0.79)
Moral conviction					
Law	.614	3.93 (1.00)	3.96 (0.94)	3.88 (1.00)	3.97 (1.03)
Justice	.583	3.39 (1.27)	3.47 (1.18)	3.39 (1.24)	3.31 (1.31)
Masks	.571	3.23 (1.40)	3.16 (1.29)	3.42 (1.31)	3.36 (1.45)
Perceived dyadic harm					
Law	.515	3.82 (1.12)	3.72 (1.14)	3.65 (1.24)	3.85 (1.22)
Justice	.527	3.26 (1.31)	3.38 (1.30)	3.23 (1.32)	3.10 (1.33)
Masks	.635	2.87 (1.44)	2.85 (1.30)	3.24 (1.32)	3.07 (1.31)
Negative moral emotion					
Law	.562	3.77 (1.18)	3.51 (1.19)	3.54 (1.30)	3.72 (1.16)
Justice	.525	3.18 (1.35)	3.16 (1.32)	3.01 (1.33)	2.97 (1.41)
Masks	.687	3.19 (1.44)	3.14 (1.35)	3.22 (1.34)	3.32 (1.34)

Note. Sample sizes vary by time point (N_{T1} = 186; N_{T2} = 165; N_{T3} = 142; N_{T4} = 126) and topic: Law (N_{T1} = 182; N_{T2} = 162; N_{T3} = 141; N_{T4} = 124), Justice (N_{T1} = 178; N_{T2} = 158; N_{T3} = 135; N_{T4} = 122), Masks (N_{T1} = 70; N_{T2} = 66; N_{T3} = 59; N_{T4} = 54). *ICC* = intraclass correlation coefficient; T = Time.

by their political opponents (that is, Biden [Trump] and his supporters). These messages served as reminders for relevant outgroup expressions in the near past (i.e., in between two measurement points), which ensured that all participants were thinking about the same relevant events when reporting on their perceptions of dyadic harm and negative emotions toward the outgroup. This standardized approach enabled a meaningful comparison between individuals' time-specific scores on the relevant variables (i.e., perceived dyadic harm and negative moral emotion), both between- and within-individuals, by reducing the potential influences of fluctuations in individuals' situational state of mind at the time of measurement. Thus, the inclusion of outgroup expressions was a methodological choice intended to reduce error variance as well as to facilitate the meaningful operationalization of both moralization drivers.

Building on the previous research (D'Amore et al., 2022), each news message was formulated to contain three core elements: (a) a description of the ingroup's general position on the relevant political topic ("Most Biden supporters are in favor of [specific topic] as a means to"), (b) a description of the outgroup's general position on this specific topic ("In contrast, most Trump supporters oppose [specific topic], as often voiced by Donald Trump himself. This can be seen in numerous ways."), and (c) two or three examples of situational expressions by the outgroup that reflect potential dyadic harm in relation to the specific topic (e.g., the *Climate* topic: "For example, instead of stimulating green energy, 'Trump rushes to finalize drilling rights in Alaska's pristine wilderness before Biden's presidential inauguration, setting up a possible auction for interested companies by the end of 2020'"). The Complete Stimulus Materials File on Open Science Framework provides a complete overview of all news messages used in the present study by wave (pp. 6–12).

Measures

The measurements reported below were scored on a 5-point scale indicating agreement (from 1 = *not at all* to 5 = *very much*), unless indicated otherwise.

Attitude Moralization

For each of the included topics (i.e., *Mask-Wearing, Paris Climate Accord, Local Law Enforcement, Supreme Court Justice Confirmation*), participants indicated the extent to which they held their attitude as a moral conviction (Skitka & Morgan, 2014; four items, e.g., "To what extent is your position on the topic of … a reflection of your core moral beliefs and convictions"; all[10] $\alpha > .92$). Attitude moralization over time was operationalized as within-subject change in an individual's topic-specific moral conviction.

Perceived Polarization

Perceived polarization was measured on a four-item scale (Koudenburg & Kashima, 2022; subscale for perceived structural polarization; "Within the U.S. society, different groups emerge based on their different views on relevant political topics", "Biden supporters and Trump supporters are in direct opposition of each other with regard to their views on relevant political topics", "In the United States, the views of Trump supporters and Biden supporters on relevant political topics are not only divided but also entrenched", and "Discussions between Trump supporters and Biden supporters on relevant political topics become more and more heated"; all $\alpha > .76$).

Network Political Homogeneity

Participants were provided with a description of network members as people with whom they, over the last month (i.e., in between two measurement waves), had discussed societal or political topics, events that happened in society, or topics related to the presidential election[11] (measure adapted from Levitan & Visser, 2008). Participants listed five network members using their first name. Next, they were asked to indicate for each network member

[10] See Supplemental Materials Table B1 (p. 5), for an overview of all topic-specific reliability coefficients by time point and sample for the model's core variables.

[11] For the precise phrasing of this description, see Complete Stimulus Materials File (pp. 6–12) on Open Science Framework.

whether and how strongly they support either Trump or Biden (or neither) within the context of the 2020 presidential election on a 5-point scale (1 = *strong Biden supporter*, 2 = *Biden supporter*, 3 = *neither Biden nor Trump supporter*, 4 = *Trump supporter*, 5 = *strong Trump supporter*). Network political homogeneity was operationalized as the average score of the five network members (i.e., ranging from 1 to 5), and reverse-coded in the Biden sample, such that higher scores represent greater network homogeneity in both samples.

Perceived Dyadic Harm

Participants indicated their perceptions of outgroup *dyadic harm* in response to the topic-specific expressions, "With these examples in mind (along with any similar examples that come to mind), please indicate whether and to what extent you agree with the following statements," on a three-item measure that taps into the three core elements of dyadic harm (D'Amore et al., 2022; Schein & Gray, 2018). Namely, the anticipated suffering of victims ("In the context of the issue of [issue], I think that Trump [Biden] and his supporters are making other people feel hurt and/or suffer from harm"), perceived intention to harm ("In the context of the issue of [issue], I think that Trump [Biden] and his supporters have the intention to harm or hurt other people"), and dyadic causality ("In the context of the issue of [issue], I think that Trump [Biden] and his supporters are directly responsible for any harm or suffering that other people could experience"; all $\alpha > .79$).

Negative Moral Emotion

Participants reported the degree to which they responded with specific emotions (i.e., *anger*, *contempt*, and *disgust*; e.g., D'Amore et al., 2022; Rozin, 1999) to the topic-specific outgroup expressions (all $\alpha > .86$). The phrasing was as follows: "Please think back once again to the behavior, intentions, and consequences of Trump [Biden] and his supporters' opposition against a national mandate for mask-wearing [another topic] in public. To what extent did they make you feel [emotion]?"

Background Measures

In addition to the above measures, participants were asked to indicate the degree to which they support presidential candidates Donald Trump and Joe Biden in the context of the current 2020 election (5-point scale, see the Procedure section above), and to report their attitude for each specific topic (5-point scale, from 1 = *strongly oppose* to 5 = *strongly support*). Finally, we included demographical measures as descriptive information of the sample in the first wave, namely age, gender, education level, and political orientation (9-point scale: 1 = *strong liberal*, to 5 = *moderate*, to 9 = *strong conservative*).

Results

Preliminary Data Inspection

Tables 1 and 2 provide means, standard deviations, and intraclass correlations for the five key variables in Sample 1 and Sample 2, respectively. In the Supplemental Materials (pp. 11–12), we elaborate on the most relevant background information to assess the (a) suitability of the context and (b) multilevel structure in the data—and we report the main conclusions below. First, the means and standard deviations of the five key variables in the model generally are in line with our expectations and hence support the suitability of the selected context for the present examination (e.g., relatively strong perceptions of polarization across time points, sufficient variance). Second, the intraclass correlations suggest that there is substantial variance at both the within- and between-subject level for all variables in both samples. This suggests that multilevel modeling is indeed required to test our hypotheses about purely within-subject processes (Hypothesis 1–Hypothesis 2) in a sensible way while accounting for stable between-subject differences on each variable.

Hypothesis Testing

What Drives Moralization Within Persons (H1.1–H1.2)?

To test whether subjective increases in perceived dyadic harm (H1.1) and negative moral emotion (H1.2) serve as within-person drivers of moralization, we used multilevel modeling because it enables the disentangling of within- from between-subject parts of time-varying predictors through specific centering approaches (e.g., Snijders & Bosker, 2011)—thereby providing a suited statistical analysis to test our theoretical predictions at the within-subject level while controlling for stable between-subject differences.[12] Specifically, to decompose the predictor variables at two levels, we used each individual's personal baseline scores (T1) to capture the time-invariant between-subject differences in each predictor at Level 2, and we centered the time-varying predictor scores around each individual's own baseline score (i.e., score on observation $T_x - T_1$) to capture the within-subject effects of the predictors at Level 1 (Hoffman, 2015, p. 397; Wang & Maxwell, 2015). Each multilevel model included a predictor's within-subject variable and controlled for its between-subject variable, using the *nlme* Package in RStudio (v3.1-155; Pinheiro et al., 2022). Following this statistical approach (see Supplemental Materials, pp. 13–14, for the detailed multilevel model specification), we tested H1.1 and H1.2 for each of the three topics in both samples.[13] As an indication of effect sizes, we report the standardized regression coefficient (*B*) and also Cohen's f^2 for each term[14] (Lorah, 2018).

Table 3 shows the results of these analyses for Biden and Trump supporters (Samples 1 and 2). In line with Hypotheses H1.1 and H1.2, the within-person effects of perceived dyadic harm and negative moral emotion are positive and significant across all topics and samples. This

[12] For further information on the importance of disentangling within- from between-subject associations in the study of political beliefs, see, for example, Brandt and Morgan (2022).

[13] We used separate models for each topic in each sample to evaluate the replicability of the findings across them, rather than combining the topics, the samples, or both in a single analysis. The most important advantage of this approach is that it does not require/assume similarity between the topics or samples in terms of the topic-specific outgroup expressions, measurements, variable means, or trends over time. Put differently, this approach evaluates generalizability, rather than assume it.

[14] Cohen's f^2 was calculated based on the R^2 for each term in the model (which was formed from pooling the intercept and residual variances; Snijders & Bosker, 1999), and should be interpreted with caution due to potential bias resulting from the multilevel modeling approach taken here (see Rights & Cole, 2018, for a discussion on the complexity of effect size measures in multilevel models and a solution restricted to mean-centered variables). Therefore, we report but do not further interpret this effect size measure.

Table 3
Direct Effects on Attitude Moralization

	Within-subject effect (T-T_1 deviation)						Between-subject effect (T_1 personal baseline)						
Predictor	*b*	*DF*	*p*	95% CI	*B*	f^2	*b*	*DF*	*p*	95% CI	*B*	f^2	R^2_{total}
Perceived dyadic harm													
S1. Climate	0.32***	509	<.001	[0.26, 0.38]	0.26	.137	0.28***	342	<.001	[0.20, 0.35]	0.27	.177	.688
S1. Justice	0.31***	463	<.001	[0.23, 0.38]	0.24	.103	0.28***	310	<.001	[0.19, 0.37]	0.27	.133	.633
S1. Masks	0.30***	510	<.001	[0.23, 0.37]	0.26	.114	0.19***	344	<.001	[0.11, 0.27]	0.20	.112	.663
S2. Law	0.26***	244	<.001	[0.19, 0.33]	0.31	.155	0.18**	179	.002	[0.07, 0.29]	0.20	.156	.681
S2. Justice	0.23***	236	<.001	[0.14, 0.31]	0.24	.088	0.22***	175	<.001	[0.10, 0.33]	0.23	.097	.541
S2. Masks	0.37***	108	<.001	[0.22, 0.52]	0.33	.221	0.28**	67	.002	[0.11, 0.44]	0.30	.241	.589
Negative moral emotion													
S1. Climate	0.42***	509	<.001	[0.36, 0.46]	0.34	.258	0.35***	342	<.001	[0.18, 0.28]	0.35	.311	.698
S1. Justice	0.49***	463	<.001	[0.42, 0.57]	0.37	.239	0.41***	310	<.001	[0.32, 0.50]	0.39	.294	.657
S1. Masks	0.31***	510	<.001	[0.25, 0.36]	0.26	.104	0.28***	344	<.001	[0.12, 0.24]	0.27	.133	.693
S2. Law	0.22***	244	<.001	[0.15, 0.29]	0.24	.086	0.08	179	.141	[−0.03, 0.08]	0.09	.087	.702
S2. Justice	0.26***	236	<.001	[0.18, 0.34]	0.28	.132	0.21***	175	<.001	[0.10, 0.31]	0.23	.136	.501
S2. Masks	0.26**	108	.001	[0.10, 0.41]	0.20	.070	0.33**	67	.001	[0.14, 0.52]	0.35	.139	.606

Note. S1 refers to Biden sample, S2 refers to Trump sample. *DF* = degrees of freedom; CI = confidence interval; T = Time.
** $p < .010$. *** $p < .001$.

suggests that subjective increases in perceived dyadic harm and negative moral emotion over time (relative to individuals' personal baseline) drive subjective increases in moralization within-individuals.[15] Thus, when individuals come to respond stronger to outgroup expressions in the news than they did before, they tend to moralize their attitudes more strongly than they did before. Furthermore, the positive and significant between-subject effects replicate previous experimental findings (D'Amore et al., 2022), suggesting that between-person differences in absolute ratings of perceived dyadic harm and negative moral emotion predict between-person differences in moralization at baseline. In sum, while the between-subject effects support previous findings, the within-subject findings extend this by highlighting the role of subjective changes in these psychological responses to outgroup expressions for explaining the subjective development of attitude moralization within-individuals.

What Predicts Drivers of Moralization Within Persons (H2.1–H2.2)?

Hypothesis H2 proposes that subjective increases in perceived polarization (H2.1) and network homogeneity (H2.2) serve as predictors of the two moralization drivers: perceived dyadic harm and negative moral emotion. To test those hypotheses, we first tested the direct effects of the two polarization variables on perceived dyadic harm, negative moral emotion, and attitude moralization using the same statistical approach as described for Hypothesis 1 above. Table 4 presents the results of the within-person effects from the multilevel regression analyses.

For perceived polarization, the effects are positive and significant on all three outcome variables across topics and samples, thus supporting H2.1. Regarding network political homogeneity, the effects are generally positive and significant across topics among Biden supporters, supporting H2.2. For Trump supporters, the effects of network homogeneity are positive and significant in some cases, and not significant in other cases. Yet, there is no indication of a meaningful pattern. Thus, overall, the results suggest that when people perceive stronger polarization and network political homogeneity than before, they also respond more strongly to the news messages in terms of stronger negative moral emotion and stronger perceived dyadic harm.

To further examine H2.1 and H2.2, we tested the indirect effects of perceived polarization and network political homogeneity on moralization via perceived dyadic harm and negative moral emotion within subjects. Specifically, we entered the relevant predictor and driver variables (both at Level 1 and at Level 2) simultaneously in the multilevel model (see Aim 1) and conducted Bootstrapped tests of average causal mediation effects at the within-subject level using the *Mediation* package in R (Tingley et al., 2014) for each topic in each sample. Table 5 displays the results of these analyses. In line with H2.1 and H2.2, most indirect effects were positive and significant. The only exception was for the Law topic among Trump supporters, where the indirect effect from network homogeneity via negative moral emotion was positive but not significant. Taken together, these results suggest that perceived polarization and network homogeneity can predict the strengthening of individuals' subjective responses to outgroup expressions in the news, and in turn, strengthen individuals' moralization of specific topics.[16]

Exploring Feedback Loops of Dynamic Change Over Time

To explore dynamic changes and reciprocal influences between polarization and moralization over time, we estimated the effect of

[15] When entering both predictors simultaneously in the model, the within-person effects reported here remain significant and generally similar in size (as would be expected from the weak correlation observed between the two predictors).

[16] We also explored parallel mediation models (Supplemental Materials Figures C1 and C2, pp. 15–16) and found support for parallel mediation via perceived dyadic harm and negative moral emotion. However, these findings should be interpreted with caution due to the moderate-to-strong bivariate correlations between the two mediators across all topics (all $r > .59$, $p < .001$).

Table 4
Within-Subject Effects of Perceived Polarization and Network Homogeneity

Predictor	Perceived dyadic harm *b* [95% CI]	Negative moral emotion *b* [95% CI]	Moralization *b* [95% CI]
Perceived polarization			
S1. Climate	0.22*** [0.10, 0.34]	0.29*** [0.17, 0.40]	0.20** [0.08, 0.31]
S1. Justice	0.28*** [0.15, 0.40]	0.23*** [0.12, 0.35]	0.22** [0.08, 0.35]
S1. Masks	0.26*** [0.16, 0.35]	0.25*** [0.16, 0.34]	0.20*** [0.10, 0.30]
S2. Law	0.53*** [0.35, 0.71]	0.30** [0.12, 0.49]	0.20** [0.06, 0.34]
S2. Justice	0.65*** [0.45, 0.84]	0.33** [0.13, 0.54]	0.16^ [−0.01, 0.30]
S2. Masks	0.22 [−0.06, 0.50]	0.27^ [0.00, 0.53]	0.02 [−0.27, 0.31]
Network homogeneity			
S1. Climate	0.21** [0.08, 0.33]	0.20*** [0.08, 0.32]	0.19** [0.06, 0.31]
S1. Justice	0.10 [−0.04, 0.23]	0.13* [0.00, 0.26]	0.26*** [0.11, 0.41]
S1. Masks	0.15* [0.04, 0.25]	0.11* [0.01, 0.21]	0.17** [0.06, 0.28]
S2. Law	0.22* [0.03, 0.42]	0.12 [−0.07, 0.32]	0.08 [−0.06, 0.23]
S2. Justice	0.16 [−0.07, 0.38]	0.29* [0.06, 0.52]	0.24* [0.04, 0.44]
S2. Masks	−0.03 [−0.36, 0.31]	0.33* [0.02, 0.64]	−0.19 [−0.54, 0.16]

Note. S1 refers to Biden sample, S2 refers to Trump sample. All effects are within-person effects. CI = confidence interval.
* $p < .050$. ** $p < .010$. *** $p < .001$. ^ $p < .100$.

perceived polarization (and network homogeneity) on moralization at the subsequent time point, as well as the reversed effect in a bidirectional model. For each topic and sample, we did so by estimating cross-lagged autoregressive structural equation models in R (Lavaan package; Rosseel, 2012). Figures 2 and 3 visualize the full cross-lagged model diagrams for perceived polarization and network homogeneity, respectively. For all topics in both samples, perceived polarization (Figure 2) and network homogeneity (Figure 3) significantly and positively predicted changes in participants' moralization over time. This suggests that the level of perceived polarization and network homogeneity at a previous time point predicts the strength of individuals' moralization later in time. This emphasizes once more the important role of polarization perceptions in predicting the development of moralization over time.

Importantly, the results also demonstrate that moralization significantly and positively predicted changes in participants' perceived polarization over time (Figure 2). Thus, this points at a reciprocal relationship in which perceived polarization and moralization reinforce one another dynamically over time. For network homogeneity, this reciprocal relationship over time is also observed among Biden supporters (Figure 3, upper diagram). Among Trump supporters, this relationship was found to be unidirectional: Network homogeneity significantly predicted individual change in moralization at the subsequent time point, but moral conviction did not significantly predict changes in network homogeneity over time (Figure 3, lower diagram). Given the explorative nature of these analyses, the results in this section should be interpreted with caution.

Finally, we explored reciprocal relationships over time between the two mediator variables and (a) attitude moralization, (b) perceived polarization, and (c) network homogeneity, using the same statistical approach. The main results of these analyses, namely the effects of each variable on each outcome at the subsequent time point, are summarized in Table 6 (Biden sample) and Table 7 (Trump sample). Overall, these results suggest that, again with the exception of network homogeneity in the Trump sample, all of the explored relationships were supported and found to be bidirectional.

More specifically, with the exception of the Law topic in the Trump sample, results suggest that both dyadic harm and negative

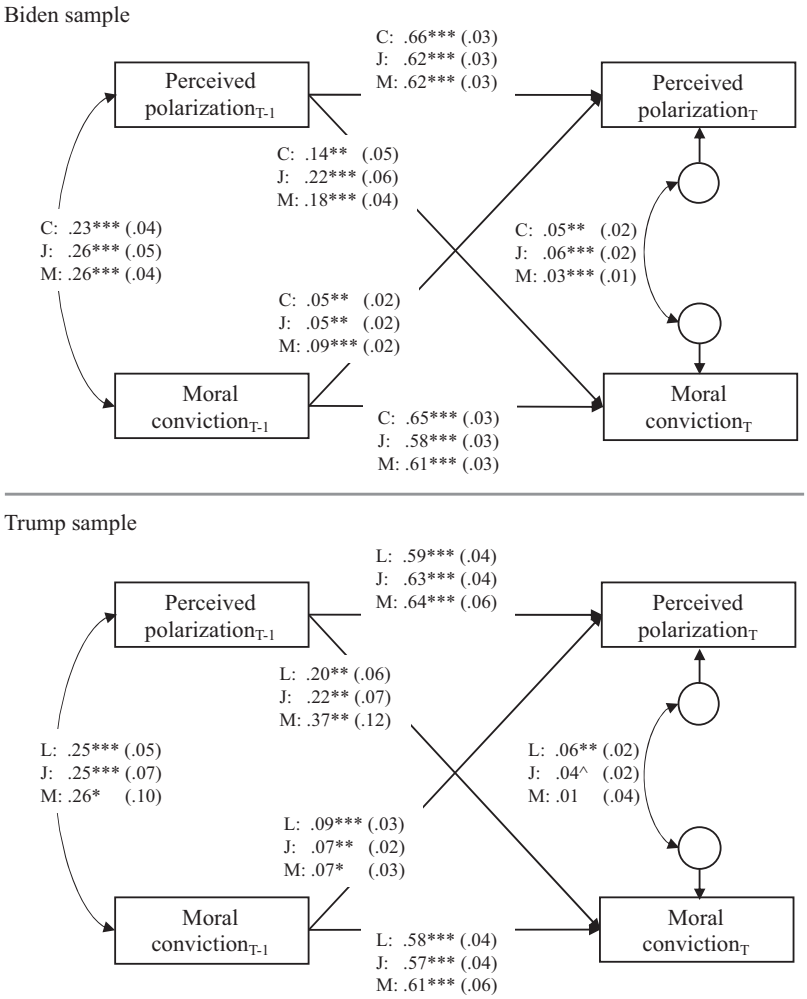
Table 5
Indirect Within-Subject Effects via Perceived Dyadic Harm and Negative Moral Emotion on Moralization

Predictor	Estimate [95% CI]: Bootstrapped mediation effect			
	Perceived dyadic harm as mediator		Negative moral emotion as mediator	
	Sample 1 (Biden)	Sample 2 (Trump)	Sample 1 (Biden)	Sample 2 (Trump)
Perceived polarization				
Climate/Law	0.07* [0.03, 0.11]	0.13* [0.09, 0.18]	0.12* [0.07, 0.18]	0.06* [0.03, 0.09]
Justice	0.08* [0.05, 0.12]	0.14* [0.08, 0.22]	0.11* [0.05, 0.16]	0.08* [0.03, 0.13]
Masks	0.07* [0.05, 0.10]	0.09^ [−0.01, 0.19]	0.07* [0.04, 0.11]	0.06* [0.00, 0.16]
Network homogeneity				
Climate/Law	0.06* [0.03, 0.11]	0.06* [0.01, 0.11]	0.08* [0.02, 0.13]	0.03 [−0.00, 0.08]
Justice	0.03^ [−0.00, 0.08]	0.04^ [−0.00, 0.10]	0.08* [0.02, 0.15]	0.07* [0.02, 0.14]
Masks	0.04* [0.02, 0.08]	−0.02 [−0.13, 0.13]	0.03* [0.01, 0.06]	0.09^ [−0.00, 0.19]

Note. CI = confidence interval.
* $p < .050$. ^ $p < .100$.

Figure 2

Cross-Lagged Longitudinal Model for Perceived Polarization and Moral Conviction



Note. Letters represent the topic: C = *Climate*, J = *Justice*, M = *Masks*, L = *Law*. Numbers represent unstandardized path estimates (*SE*). *SE* = standard error.
$^{*}p < .050$. $^{**}p < .010$. $^{***}p < .001$. $^{\wedge}p < .100$.

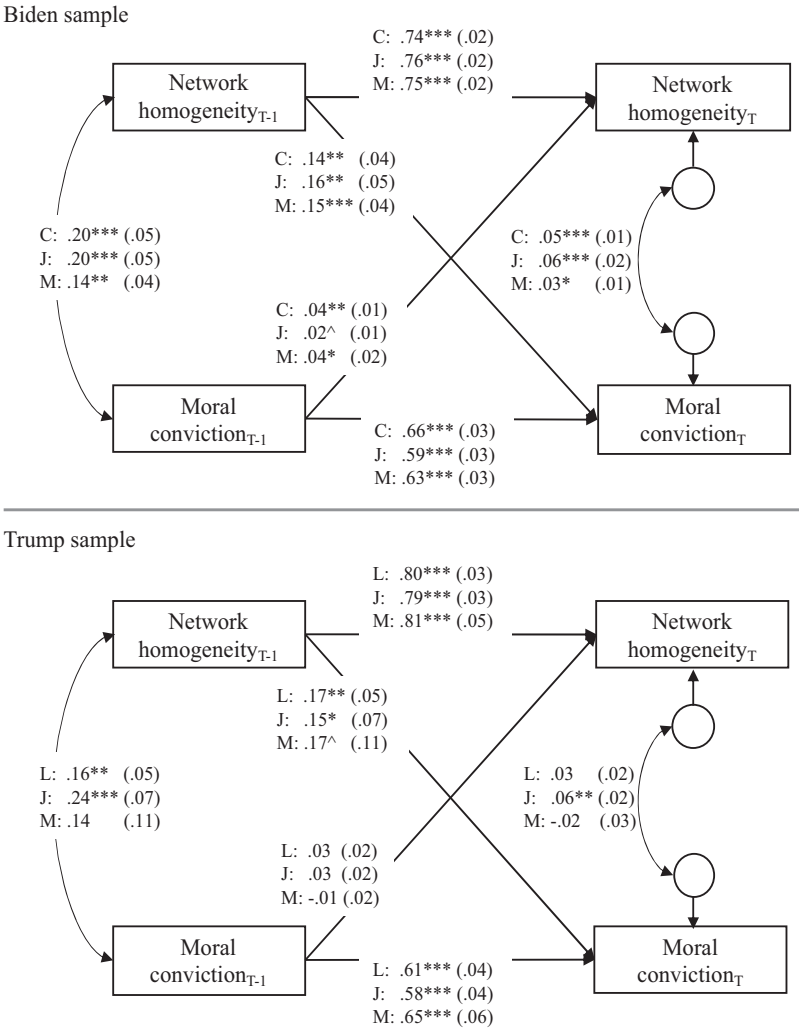
moral emotion are significant and positive predictors of moralization at the subsequent time point in both samples. Considering the reversed effects, the results also indicate a (marginally) significant effect of moral conviction on both perceived dyadic harm and negative moral emotion at the subsequent time point, with the exception of negative moral emotion for the Justice topic in the Trump sample. Similarly, looking at the dynamic relationships between perceived polarization and both perceived dyadic harm and negative moral emotion, both directional effects are significant for all topics and samples, again suggesting dynamic reciprocal relationships over time. For network homogeneity, however, such dynamic reciprocal relationships with perceived dyadic harm and negative moral emotion are only found among Biden supporters (for all topics) but not among Trump supporters. Taken together, these findings point at bidirectional relationships between each of the two mediators (perceived dyadic harm and negative moral emotion) and perceived polarization as well as attitude moralization, each reinforcing one another dynamically over time. Although exploratory in nature, these findings are therefore in line with the idea that there are positive feedback loops between the proposed predictors, drivers, and moralization over time, suggesting an upward spiral of change over time.

General Discussion

The findings of this longitudinal research in the context of the U.S. 2020 election campaign supported our integrative model, explaining when and how individuals strengthen their moral convictions on specific topics over time while being embedded in a dynamically polarized context. Moreover, the findings support the potential for expanding the model to include positive feedback loops between perceived polarization and attitude moralization

Figure 3

Cross-Lagged Longitudinal Model for Network Homogeneity and Moral Conviction



Note. Letters represent the topic: C = *Climate*; J = *Justice*; M = *Masks*; L = *Law*. Numbers represent unstandardized path estimates (*SE*). *SE* = standard error.
* $p < .050$. ** $p < .010$. *** $p < .001$. ^ $p < .100$.

over time. A better understanding of how attitude moralization comes about is of great importance because individuals' moral convictions are powerful motivators of political engagement and activism (e.g., Agostini & van Zomeren, 2021) yet also foster increasingly destructive, zero-sum forms of political conflict within polarized contexts (e.g., Finkel et al., 2020). As little is known about what makes individuals develop strong moral convictions over time within dynamic and polarized contexts in the first place, the present findings are the first to test and support our conceptual model. We move beyond previous experimental research (D'Amore et al., 2022) by uniquely examining the within-person drivers and predictors of attitude moralization in polarized contexts over time.

Specifically, we found support for our model's predictions that an individual's increase in moralization occurs when they perceive stronger dyadic harm (H1.1) and respond with stronger negative moral emotion (H1.2) to potentially harmful outgroup expressions in the news than they did before. We also found support for the model's predictions that subjective increases in individuals' perceptions of polarization (H2.1) and network political homogeneity (H2.2) indirectly strengthen individual moralization tendencies over time by strengthening the degree to which they perceive dyadic harm and respond with negative moral emotion to those expressions relative to before. Finally, moving beyond the within-person moralization process, we explored and found positive feedback loops between perceived polarization and attitude moralization over time (H3), which suggests that they dynamically reinforce each other over time.

It is important to note that results generally replicated across different specific topics and across samples of Biden and Trump

Table 6
Overview of All Cross-Lagged Longitudinal Path Estimates (Biden Supporters)

Predictor	$^{\Delta}$Perceived polarization $Time_{T-1} \rightarrow Time_T$	$^{\Delta}$Network homogeneity $Time_{T-1} \rightarrow Time_T$	$^{\Delta}$Perceived dyadic harm $Time_{T-1} \rightarrow Time_T$	$^{\Delta}$Negative moral emotion $Time_{T-1} \rightarrow Time_T$	$^{\Delta}$Moral conviction $Time_{T-1} \rightarrow Time_T$
Perceived polarization$_{T-1}$					
Climate			0.23 (0.05)***	0.19 (0.05)***	0.14 (0.05)**
Justice			0.20 (0.05)***	0.18 (0.05)***	0.22 (0.06)***
Mask			0.24 (0.04)***	0.19 (0.04)***	0.18 (0.04)***
Network homogeneity$_{T-1}$					
Climate			0.20 (0.04)***	0.18 (0.04)***	0.14 (0.04)**
Justice			0.12 (0.05)**	0.12 (0.04)**	0.16 (0.05)**
Mask			0.14 (0.04)***	0.11 (0.04)**	0.15 (0.04)***
Perceived dyadic harm$_{T-1}$					
Climate	0.06 (0.02)***	0.04 (0.01)**			0.10 (0.03)***
Justice	0.04 (0.02)**	0.03 (0.01)*			0.18 (0.03)***
Mask	0.08 (0.02)***	0.05 (0.02)**			0.05 (0.03)^
Negative moral emotion$_{T-1}$					
Climate	0.07 (0.02)***	0.06 (0.01)***			0.17 (0.03)***
Justice	0.05 (0.02)**	0.04 (0.01)**			0.22 (0.04)***
Mask	0.10 (0.02)***	0.05 (0.02)**			0.12 (0.03)***
Moral conviction$_{T-1}$					
Climate	0.05 (0.02)**	0.04 (0.01)**	0.08 (0.03)*	0.09 (0.03)**	
Justice	0.05 (0.02)**	0.02 (0.01)^	0.10 (0.03)**	0.08 (0.03)**	
Mask	0.09 (0.02)***	0.04 (0.02)*	0.05 (0.03)^	0.07 (0.03)**	

Note. Unstandardized path estimates (standard error). T = Time.
* $p < .050$. ** $p < .010$. *** $p < .001$. ^ $p < .100$.

supporters: The more individuals strengthened their perceptions of polarization and network homogeneity from one time to another, the more they strengthened their perceptions of dyadic harm and negative moral emotions in response to reminders of expressions of the political outgroup, which in turn lead to increases in their moralization on specific topics from one time point to another. Moreover, the explorative analyses identified scope for expansion of the model by identifying that perceived polarization and attitude moralization mutually reinforced each other over time. Below, we outline the implications and limitations of our findings, as well as directions for future research and theorizing on structural polarization and attitude moralization.

Table 7
Overview of All Cross-Lagged Longitudinal Path Estimates (Trump Supporters)

Predictor	$^{\Delta}$Perceived polarization $Time_{T-1} \rightarrow Time_T$	$^{\Delta}$Network homogeneity $Time_{T-1} \rightarrow Time_T$	$^{\Delta}$Perceived dyadic harm $Time_{T-1} \rightarrow Time_T$	$^{\Delta}$Negative moral emotion $Time_{T-1} \rightarrow Time_T$	$^{\Delta}$Moral conviction $Time_{T-1} \rightarrow Time_T$
Perceived polarization$_{T-1}$					
Law			0.21 (0.08)**	0.21 (0.08)**	0.20 (0.06)**
Justice			0.41 (0.08)***	0.42 (0.09)***	0.22 (0.07)**
Mask			0.41 (0.12)***	0.29 (0.12)*	0.37 (0.12)**
Network homogeneity$_{T-1}$					
Law			0.26 (0.06)***	0.09 (0.07)	0.17 (0.05)**
Justice			0.30 (0.08)***	0.06 (0.08)	0.15 (0.07)*
Mask			0.14 (0.11)	0.10 (0.10)	0.17 (0.11)^
Perceived dyadic harm$_{T-1}$					
Law	0.09 (0.02)***	0.01 (0.02)			0.04 (0.04)
Justice	0.06 (0.02)**	0.01 (0.02)			0.19 (0.05)***
Mask	0.05 (0.03)^	0.03 (0.03)			0.12 (0.06)*
Negative moral emotion$_{T-1}$					
Law	0.06 (0.02)**	0.01 (0.02)			0.02 (0.03)
Justice	0.09 (0.02)***	0.00 (0.02)			0.11 (0.04)**
Mask	0.09 (0.03)**	−0.01 (0.02)			0.15 (0.06)*
Moral conviction$_{T-1}$					
Law	0.09 (0.03)***	0.03 (0.02)	0.11 (0.06)^	0.13 (0.05)*	
Justice	0.07 (0.02)**	0.03 (0.02)	0.08 (0.04)*	0.07 (0.05)	
Mask	0.07 (0.03)*	−0.01 (0.02)	0.07 (0.06)	0.06 (0.06)	

Note. Unstandardized path estimates (standard error). T = Time.
* $p < .050$. ** $p < .010$. *** $p < .001$. ^ $p < .100$.

Implications

This first longitudinal test of our integrative model moves beyond previous work on attitude moralization to explain and test the *within-person process* of attitude moralization over time via perceptions of polarization. Furthermore, by connecting the psychological process of moralization over time with a dynamic and structurally polarized context, the present findings extend existing research and theory on moralization and polarization in at least six ways.

First, previous work laid the groundwork for understanding when and how attitude moralization develops psychologically (Brandt et al., 2015; D'Amore et al., 2022; Leal et al., 2023; Pauls et al., 2022; Wisneski & Skitka, 2017), but this work did not provide insight into how the broader social and societal context may contribute to this psychological process over time. In contrast, the present study helps to better understand how perceived structural divisions between two political groups in society give rise to attitude moralization through a similar process among people from both sides of the political divide. As such, these findings newly suggest that the broader context of polarization plays an important role in the process of attitude moralization across specific political topics. Specifically, our study showed that increasing perceptions of structural polarization over time can strengthen individuals' psychological response to potentially harmful expressions in the news by one and the same outgroup, and in this way predict the subjective strengthening of moralization—indeed, we found evidence in support of this process across multiple different topics and within each of the samples we collected. This illustrates the importance of studying attitude moralization *in context*, and underscores the potential of groups as agents of conflict, change, and moralization.

Second, our model and findings offer a *socially grounded* perspective on understanding what the underlying motivational drivers for attitude moralization within these contexts may be. This is different from other studies that conceptualized predictors of moralization as *intrinsic* to the topic of moralization (e.g., Feinberg et al., 2019; Rozin, 1999; Wisneski & Skitka, 2017), implying that the motivation for moralization is rooted in changing feelings and cognitions about the concrete topic at hand.[17] In this view, moralization can thus be understood as emerging from a changed understanding of the specific topic, reflecting a proactive motive to stay true to one's fundamental moral principles. We break new ground by highlighting a socially grounded starting point through which moralization may develop within polarized contexts. Our findings demonstrate that the perception of polarization in society and one's network can, even without reference to any specific topic, strengthen one's tendency to moralize a variety of political topics, through strengthening one's value-protective responses to a specific political outgroup and their expressions.

This insight is particularly important because it implies that individual attitude moralization about specific topics may be driven by a broader motivation to protect ingroup values from outgroup attacks, enabled by association processes within polarized contexts (van Zomeren et al., 2023). Specifically, when increasing perceptions of structural polarization are combined with repeated exposure to potentially harmful outgroup expressions related to a specific topic, this strengthens the subjective association between the topic at hand and one's relevant core values (i.e., a phenomenon called moral piggybacking; Feinberg et al., 2019; Rozin, 1999)—and hence motivate individuals to moralize these attitudes to protect those core values from potential attacks by the outgroup. Attitude moralization within the individual thus seems to be enabled and predicted by (perceived) aspects of the social context that trigger the motivation to protect ingroup values from perceived outgroup attack. Hence, although moralization within-individuals is by definition a psychological process, what predicts it within structurally polarized contexts can thus be something fundamentally social.

Third, our perspective implies that it is important to differentiate between forms of polarization, and perhaps even beyond Koudenburg and Kashima's (2022) distinction that we used in this article. If attitude moralization reflects a value-protective response to perceived harmful outgroup expressions, then the structural dimension of perceived polarization should be key for perceptions of polarization to strengthen moralization via such value-protective processes. This fits with recent insights on the consequences of perceiving *moral polarization* between partisan groups (i.e., structural value-based conflict), where the perception of moral polarization indirectly predicted increased support for strong leaders who are willing to fight to protect the core values of their political ingroup (Crimston et al., 2022). In contrast, without focus on the structural dimension of group divisions, our theoretical rationale would imply that perceptions of opinion polarization in society are unlikely to predict value-protective responses to situational outgroup expressions. Supporting this line of thought, research suggests that when perceptions of polarization focus on the magnitude of opinion differences in society, this can even decrease feelings of threat and negative moral emotions toward attitudinal opponents (Koudenburg & Kashima, 2022). Future research can further examine which forms of perceived polarization are most likely to strengthen attitude moralization within persons.

Fourth, our findings suggest that perceived polarization may "facilitate" the transformation of relatively nonmoral political disagreement into moralized, zero-sum conflict between groups in society over time. Even though specific political topics are sometimes considered to be intrinsically connected to morality (e.g., on theoretical grounds), the literature on moral conviction has consistently demonstrated much variation in the extent to which individuals themselves consider specific topics as morally relevant (e.g., Ryan, 2014). As such, political disagreement between groups can have a relatively nonmoralized or moralized meaning. While nonmoralized disagreement is usually nonproblematic and relatively easy to resolve, such as through compromise, moral conflict tends to be difficult to resolve, fuels the escalation of conflict, and pose a threat to healthy democratic functioning (e.g., Finkel et al., 2020; Kovacheff et al., 2018; Mooijman et al., 2018; Skitka et al., 2021). We theoretically reasoned and empirically demonstrated in

[17] For example, the push–pull model of moralization (Feinberg et al., 2019) proposed and found that, after confronting individuals with shocking videos about the suffering of animals in the meat industry, topic-specific factors such as the perceived suffering of animals due to human meat consumption could "push" individuals to moralize one's attitude against meat-eating, whereas other factors such as perceived sacrifices when refraining from eating meat could "pull" individuals away from moralization. In turn, moralization was explained by moral emotions toward animals and through the mechanism of *moral piggybacking*, referring to the process of suddenly connecting the specific topic (e.g., meat-eating) with one's existing moral principles (e.g., causing harm is morally wrong), hence infusing the topic at hand with moral meaning (Feinberg et al., 2019; Rozin, 1999).

this article that the perception of structural intergroup differences in society can make individuals strengthen their moralization of attitudes on specific topics (i.e., our psychological outcome variable of interest). At the societal level, this implies that greater perceptions of polarization can strengthen attitude moralization in both political groups, and thus strengthen the moral meaning of the existing intergroup differences at the large scale—that is, greater moral conflict in society.

Intriguingly, the more explorative findings of this study point to the possibility of expanding our model and thereby our understanding of the complex relationship between structural polarization and attitude moralization. Indeed, findings identified a reciprocal relationship between moralization and perceived polarization, influencing each other dynamically over time. This can be linked to research suggesting that while there are many differences between Democrats and Republicans, people tend to overestimate levels of polarization (e.g., Baldassarri & Bearman, 2007; Finkel et al., 2020; Westfall et al., 2015). Our current findings suggest that one reason why this may be the case is because of attitude moralization: Moral convictions likely amplify perceived differences between people who share one's moral beliefs and those who disagree, overshadowing actual differences. If these (mis)-perceptions of polarization predict drivers of moralization, as our findings suggest, this does not seem to bode well for the future of such polarized societies. We can see at least two different scenarios. The first is one in which polarization reinforces moralization, and moralization further reinforces polarization in a self-reinforcing cycle, leading to further societal division and eventual breakdown (cf. work on societal schisms, McCoy & Somer, 2019; Sani & Reicher, 1998, 1999).

Perhaps a second scenario may be more likely. This scenario includes de-escalating factors that we did not yet focus on (i.e., conceptualize and measure) and hence did not observe in this specific study. Nevertheless, in the natural variance in the data of the present study, for example, we note that there are interesting trends over time of mean-level moralization and polarization but also of mean-level "de-moralization" and "de-polarization," and we suspect that to predict the future of polarizing societies, we need to be able to explain both escalation and de-escalation patterns at the macro-level. Although the present study design does not enable a sensible assessment of such patterns in society or its potential causes, we can speculate that an increased salience of political conflict about a specific topic (e.g., increased media attention at the societal level and/or increased discussion frequency at the network level) encourages moralization, while decreased salience encourages de-moralization.[18] Overall, then, we do *not* claim that polarization and moralization are necessarily detrimental for society and democracy. However, when viewed through the lens of our conceptual model and its suggested feedback loop, we do see good reason to be concerned. Future research is needed to examine which factors escalate and de-escalate political and moral conflict in polarized contexts.

Finally, a practical implication of our findings is that if polarization predicts increased moralization in response to outgroup expressions in the news, our research points at two intervention points in order to buffer against this triggering of moralization. First, considering that mass media outlets can be partisan in both the selection and framing of politically relevant news messages (e.g., Druckman et al., 2019), and that the inclusion of moral-emotional content in online messages significantly amplifies its spread (Brady et al., 2020), it would not surprise us to find that online media often includes numerous cues to outgroup dyadic harm framed through a partisan lens. Indeed, the election context appeared to provide plenty of conceptually relevant cues to dyadic harm in the form of news messages, as we used this to select specific outgroup expressions as reminders of real-time events for our participants and assess their responses. While we are not the first to imply that media outlets are agents of increased perceived polarization in society (e.g., Druckman et al., 2019; Levendusky & Malhotra, 2016b), we also want to emphasize that the media, for the very same reason, can be effective agents of de-polarization as well. Second, this is not to suggest that we view individuals themselves as passive victims of partisan media reporting on upcoming elections; in fact, our findings also show that the homogeneity of one's network similarly predicts drivers of moralization. It follows that, independent of media coverage, individuals who reach out to those with different political views and hence contribute to establishing a more heterogenous network can agentically help reduce the power of potential moralization drivers amid polarized and dynamic contexts, such as a presidential election.

Limitations and Future Directions

This research has at least three limitations that future research could seek to address. First, while we tested and found empirical support for the hypotheses of our model across a various topics and two different samples, it is still an open question whether and under which conditions the present findings can be generalized across other topics, outgroups, and polarized contexts. Theoretically, being embedded in the context of structural polarization, our model's theoretical predictors for moralization are all conceptualized as conditional upon the existence of (at least) two opposing groups in society. Nevertheless, within this context of polarization, there may be boundary conditions when considering the effect of perceived polarization on individuals' psychological responses to potentially harmful outgroup expressions and attitude moralization. For example, while the present study focused on a two-party political context, one alternative approach would be to focus on a context in which polarization revolves around multiple different groups, such as political contexts with multiparty systems featuring multiple political outgroups, and test for potential boundary conditions.

Second, we acknowledge that the design of this longitudinal study is not experimental and hence we cannot causally disentangle the effects of perceived polarization and exposure to the outgroup expressions. Indeed, the chosen context of the present research was a highly polarized one across all four waves and all participants were reminded of the same outgroup expressions. For the present research, such a polarized context was suited to test the model's hypotheses as it allowed us to obtain and include a variety of real-time outgroup expressions from the in a relatively natural way (i.e., participants being faced with news messages one may have encountered in daily life). One limitation of this standardized

[18] For example, the present data suggested that in both samples, attitude moralization about the Justice topic seemed to decrease after the confirmation of the new Supreme Court Justice took place and media attention decreased. Given that this study was specifically designed to assess within-subject processes rather than macrolevel trends, however, we cannot assume that the observed trends in this data are representative for the population and hence no direct conclusions should be made on the basis of this study alone.

approach, however, is that this research could only assess whether *change* in participants' responses to the news messages predicted change in their moralization, whereas it cannot be used to identify whether participants' *exposure* to the news messages as part of the survey (vs. no exposure) triggered any change in their responses. Although this could potentially be solved by adding a "control" condition in which individuals are not reminded of such outgroup expressions, the resulting measures of value-protective responses (dyadic harm and negative moral emotions) would in this scenario not be based on the same situation as for participants who were reminded of outgroup expressions in the news. In our view, this would introduce a lack of comparability in key variables of our model, which we aimed to avoid. Similarly, due to the highly polarized context, this research could only assess whether *change* in perceptions of polarization predict change in individuals' psychological responses, whereas it could not evaluate whether perceiving polarization (vs. no polarization) is a requirement for perceived outgroup dyadic harm to function as a driver of moralization. Future research may consider designing a controlled experimental study specifically aimed at examining such potential boundary conditions.

Furthermore, although the present research was designed to examine the process of moralization in particular, it also offered intriguing exploratory insights about the interplay between moralization and perceived polarization that are important to examine further. As described earlier, the exploratory analyses suggest a mutually reinforcing relationship between moralization and perceived polarization. This raises the question of whether or not the two moralization drivers may function as mediators in both causal directions. To examine this question, future research could design a controlled pretest–posttest study in which outgroup dyadic harm (vs. control) is manipulated and both perceived polarization and moralization are measured prior and after exposure. In contrast, if the aim is to identify the exact order in which each of the model's core variables influence each other over time or how the variables may interact to produce a reinforcing spiral, a more suited approach may be microlongitudinal research (i.e., increasing the number of repeated measures with shorter time intervals).

Likewise, as would be expected from our model, our samples within this polarized context showed moderate levels of preexisting moral convictions at baseline, with fluctuating sample-means over time both toward stronger moralization as well as "de-moralization" for some specific topics. While our theoretical model and hypotheses were specifically about the process of moralization, the empirical findings could suggest at least two possibilities with regard to de-moralization: (a) the factors relevant to the process of moralization and that of de-moralization may be the same (i.e., reversed predictions), or (b) the factors studied here were only relevant to moralization and not de-moralization, such that the de-moralization that occurred in our participants added unexplained variance to our models. Empirically, the present data demonstrate similar effects for topics that either showed a mean-level trend of moralization (e.g., the Mask-Wearing topic in the Biden sample) or of de-moralization (e.g., the Justice topic in the same sample), which may suggest that similar factors are predictive of both. However, given that even less is known about de-moralization than about moralization (see Skitka et al., 2021), the present findings should not simply be interpreted as evidence for similarity between the two processes. In fact, once moralized, individuals are thought to be resistant to peer influence (e.g., Skitka & Morgan, 2014), which suggests a potentially different de-moralization process. This is clearly an important and novel direction for future research to pursue.

Furthermore, Skitka et al. (2021) theorizing suggests that the type of moralization studied in our research (*moral amplification*, which describes further moralization of existing convictions; Rhee et al., 2019) would make social factors as group loyalty more likely to be important for moralization (Skitka et al., 2021). If true, then this raises the question whether moralization of nonmoralized attitudes (i.e., *moral recognition*) will also be predicted by perceptions of polarization and network homogeneity, such as when individuals are confronted with potentially harmful expressions by the political outgroup signaling their opposition on a relatively unknown or new (and hence weakly moralized) political topic. We would expect this to be the case, but future research is needed to answer this question empirically.

Third, we selected the specific topics for moralization based on the degree to which they are discussed in the media and were assumed to reflect the political divide. For most topics, this appeared to be true, as reflected in participants' attitudes in both samples. However, the Mask-Wearing topic did not meet this assumption in the Trump sample (i.e., the majority supported a mask-wearing mandate), and indeed results deviated from the general pattern for this topic and sample, especially in terms of perceived polarization and network homogeneity. It is possible that the topic did not fit the polarized context very well, given the fact that only a small percentage of Trump supporters were against a mask-wearing mandate, which suggests that opinions were not as clearly divided between the groups. Because of this, and also because of the smaller sample size for this specific case, these results should be interpreted with caution. Future research could examine whether and how the replicability of the effects of polarization depends on characteristics of the specific topic studied at hand.

Conclusion

Does polarization predict the development of attitude moralization in the context of the U.S. 2020 elections? Our longitudinal study offered an affirmative answer to this question. We found that the perception of structural polarization at both the societal level and the network level (i.e., politically homogeneous networks) played an important role in individuals' moralization of specific attitudes over time via predicting their psychological, value-protective responses to potentially harmful expressions by the political outgroup. Specifically, our findings imply that increased perceptions of polarization over time can strengthen individuals' tendency to moralize attitudes across various political topics simultaneously, via strengthening individuals' subjective perceptions of dyadic (intentional) harm by and negative emotions against one and the same political outgroup over time. This suggests that the motivational drive for attitude moralization—set against the backdrop of potentially harmful outgroup expressions—can be understood as an essentially social one. Furthermore, the exploratory findings of this study identified a dynamic feedback loop between perceived polarization and attitude moralization, mutually influencing each other over time. This research thereby sheds light on the social contextual factors that transform relatively nonmoral disagreements between groups into moralized conflict, as observed in the U.S. 2020 elections and its aftermath.

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

Atari, M., Davani, A. M., Kogon, D., Kennedy, B., Ani Saxena, N., Anderson, I., & Dehghani, M. (2022). Morally homogeneous networks and radicalism. *Social Psychological & Personality Science*, *13*(6), 999–1009. https://doi.org/10.1177/19485506211059329

Baldassarri, D., & Bearman, P. (2007). Dynamics of political polarization. *American Sociological Review*, *72*(5), 784–811. https://doi.org/10.1177/000312240707200507

Berinsky, A. J., Huber, G. A., & Lenz, G. S. (2012). Evaluating online labor markets for experimental research: Amazon.com's Mechanical Turk. *Political Analysis*, *20*(3), 351–368. https://doi.org/10.1093/pan/mpr057

Brady, W. J., Crockett, M. J., & Van Bavel, J. J. (2020). The MAD model of moral contagion: The role of motivation, attention and design in the spread of moralized content online. *Perspectives on Psychological Science*, *15*(4), 978–1010. https://doi.org/10.1177/1745691620917336

Brandt, M. J., & Morgan, G. S. (2022). Between-person methods provide limited insight about within-person belief systems. *Journal of Personality and Social Psychology*, *123*(3), 621–635. https://doi.org/10.1037/pspp0000404

Brandt, M. J., Wisneski, D. C., & Skitka, L. J. (2015). Moralization and the 2012 U.S. presidential election campaign. *Journal of Social and Political Psychology*, *3*(2), 211–237. https://doi.org/10.5964/jspp.v3i2.434

Brown, J. R., & Enos, R. D. (2021). The measurement of partisan sorting for 180 million voters. *Nature Human Behaviour*, *5*(8), 998–1008. https://doi.org/10.1038/s41562-021-01066-z

Cole Wright, J., Cullum, J., & Schwab, N. (2008). The cognitive and affective dimensions of moral conviction: Implications for attitudinal and behavioral measures of interpersonal tolerance. *Personality and Social Psychology Bulletin*, *34*(11), 1461–1476. https://doi.org/10.1177/0146167208322557

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

D'Amore, C. (2024). *Connecting polarization with moralization*. https://doi.org/10.17605/OSF.IO/F9P4E

D'Amore, C., van Zomeren, M., & Koudenburg, N. (2022). Attitude moralization within polarized contexts: An emotional value-protective response to dyadic harm cues. *Personality and Social Psychology Bulletin*, *48*(11), 1566–1579. https://doi.org/10.1177/01461672211047375

Druckman, J. N., Gubitz, S. R., Lloyd, A. M., & Levendusky, M. S. (2019). How incivility on partisan media (de)polarizes the electorate. *The Journal of Politics*, *81*(1), 291–295. https://doi.org/10.1086/699912

Echterhoff, G., & Higgins, E. T. (2017). Creating shared reality in interpersonal and intergroup communication: The role of epistemic processes and their interplay. *European Review of Social Psychology*, *28*(1), 175–226. https://doi.org/10.1080/10463283.2017.1333315

Ellemers, N., & van den Bos, K. (2012). Morality in groups: On the social-regulatory functions of right and wrong. *Social and Personality Psychology Compass*, *6*(12), 878–889. https://doi.org/10.1111/spc3.12001

Ellemers, N., van der Toorn, J., Paunov, Y., & van Leeuwen, T. (2019). The psychology of morality: A review and analysis of empirical studies published from 1940 through 2017. *Personality and Social Psychology Review*, *23*(4), 332–366. https://doi.org/10.1177/1088868318811759

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

Feinberg, M., Kovacheff, C., Teper, R., & Inbar, Y. (2019). Understanding the process of moralization: How eating meat becomes a moral issue. *Journal of Personality and Social Psychology*, *117*(1), 50–72. https://doi.org/10.1037/pspa0000149

Finkel, E. J., Bail, C. A., Cikara, M., Ditto, P. H., Iyengar, S., Klar, S., Mason, L., McGrath, M. C., Nyhan, B., Rand, D. G., Skitka, L. J., Tucker, J. A., Van Bavel, J. J., Wang, C. S., & Druckman, J. N. (2020). Political sectarianism in America. *Science*, *370*(6516), 533–536. https://doi.org/10.1126/science.abe1715

Goldenberg, A., Abruzzo, J., Willer, R., Halperin, E., & Gross, J. (2020). *Homophily and acrophily as drivers of political segregation*. https://doi.org/10.31219/osf.io/9xetr

Gray, K., MacCormack, J. K., Henry, T., Banks, E., Schein, C., Armstrong-Carter, E., Abrams, S., & Muscatell, K. A. (2022). The affective harm account (AHA) of moral judgment: Reconciling cognition and affect, dyadic morality and disgust, harm and purity. *Journal of Personality and Social Psychology*, *123*(6), 1199–1222. https://doi.org/10.1037/pspa0000310

Haidt, J. (2012). *The righteous mind: Why good people are divided by politics and religion*. Vintage.

Hoffman, L. (2015). *Longitudinal analysis: Modeling within-person fluctuation and change*. Routledge. https://doi.org/10.4324/9781315744094

Hornsey, M. J. (2008). Social identity theory and self categorization theory: A historical review. *Social and Personality Psychology Compass*, *2*(1), 204–222. https://doi.org/10.1111/j.1751-9004.2007.00066.x

Klein, E. (2020). *Why we're polarized*. Simon and Schuster.

Koudenburg, N. (2018). Regulating shared reality with micro-dynamics in the form of conversation. *Current Opinion in Psychology*, *23*, 47–51. https://doi.org/10.1016/j.copsyc.2017.12.002

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

Kovacheff, C., Schwartz, S., Inbar, Y., & Feinberg, M. (2018). The problem with morality: Impeding progress and increasing divides. *Social Issues and Policy Review*, *12*(1), 218–257. https://doi.org/10.1111/sipr.12045

Leal, A., van Zomeren, M., Gordijn, E., Reifen Tagar, M., & Halperin, E. (2023). *Attitude moralization within intergroup contexts: When and how do immoral outgroup violations trigger moralization?* [Manuscript submitted for publication].

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., & Malhotra, N. (2016a). Does media coverage of partisan polarization affect political attitudes? *Political Communication*, *33*(2), 283–301. https://doi.org/10.1080/10584609.2015.1038455

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

Levitan, L. C., & Visser, P. S. (2008). The impact of the social context on resistance to persuasion: Effortful versus effortless responses to counter-attitudinal information. *Journal of Experimental Social Psychology*, *44*(3), 640–649. https://doi.org/10.1016/j.jesp.2007.03.004

Levitan, L. C., & Visser, P. S. (2009). Social network composition and attitude strength: Exploring the dynamics within newly formed social networks. *Journal of Experimental Social Psychology*, *45*(5), 1057–1067. https://doi.org/10.1016/j.jesp.2009.06.001

Levitan, L. C., & Wronski, J. (2014). Social context and information seeking: Examining the effects of network attitudinal composition on engagement with political information. *Political Behavior*, *36*(4), 793–816. https://doi.org/10.1007/s11109-013-9247-z

Lorah, J. (2018). Effect size measures for multilevel models: Definition, interpretation, and TIMSS example. *Large-Scale Assessments in Education*, *6*(1), Article 8. https://doi.org/10.1186/s40536-018-0061-2

Lupton, R. N., Singh, S. P., & Thornton, J. R. (2015). The moderating impact of social networks on the relationships among core values, partisanship,

and candidate evaluations. *Political Psychology*, *36*(4), 399–414. https://doi.org/10.1111/pops.12102

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

Mooijman, M., Hoover, J., Lin, Y., Ji, H., & Dehghani, M. (2018). Moralization in social networks and the emergence of violence during protests. *Nature Human Behaviour*, *2*(6), 389–396. https://doi.org/10.1038/s41562-018-0353-0

Pauls, I. L., Shuman, E., Van Zomeren, M., Saguy, T., & Halperin, E. (2022). Does crossing a moral line justify collective means? Explaining how a perceived moral violation triggers normative and nonnormative forms of collective action. *European Journal of Social Psychology*, *52*(1), 105–123. https://doi.org/10.1002/ejsp.2818

Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., & R Core Team. (2022). *nlme: Linear and nonlinear mixed effects models* [R package Version 3.1-155]. https://CRAN.R-project.org/package=nlme

Price, V., Nir, L., & Cappella, J. N. (2006). Normative and informational influences in online political discussions. *Communication Theory*, *16*(1), 47–74. https://doi.org/10.1111/j.1468-2885.2006.00005.x

Rhee, J. J., Schein, C., & Bastian, B. (2019). The what, how, and why of moralization: A review of current definitions, methods, and evidence in moralization research. *Social and Personality Psychology Compass*, *13*(12), Article e12511. https://doi.org/10.1111/spc3.12511

Rights, J. D., & Cole, D. A. (2018). Effect size measures for multilevel models in clinical child and adolescent research: New r-squared methods and recommendations. *Journal of Clinical Child and Adolescent Psychology*, *47*(6), 863–873. https://doi.org/10.1080/15374416.2018.1528550

Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, *48*(2), 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

Ryan, T. J. (2014). Reconsidering moral issues in politics. *The Journal of Politics*, *76*(2), 380–397. https://doi.org/10.1017/S0022381613001357

Ryan, T. J. (2017). No compromise: Political consequences of moralized attitudes. *American Journal of Political Science*, *61*(2), 409–423. https://doi.org/10.1111/ajps.12248

Sani, F., & Reicher, S. (1998). When consensus fails: An analysis of the schism within the Italian Communist Party. *European Journal of Social Psychology*, *28*(4), 623–645. https://doi.org/10.1002/(SICI)1099-0992(199807/08)28:4<623::AID-EJSP885>3.0.CO;2-G

Sani, F., & Reicher, S. (1999). Identity, argument and schism: Two longitudinal studies of the split in the church of England over the ordination of women to the priesthood. *Group Processes & Intergroup Relations*, *2*(3), 279–300. https://doi.org/10.1177/1368430299023005

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

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

Skitka, L. J. (2002). Do the means always justify the ends, or do the ends sometimes justify the means? A value protection model of justice reasoning. *Personality and Social Psychology Bulletin*, *28*(5), 588–597. https://doi.org/10.1177/0146167202288003

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., & Morgan, G. S. (2014). The social and political implications of moral conviction. *Political Psychology*, *35*(S1), 95–110. https://doi.org/10.1111/pops.12166

Snijders, T. A. B., & Bosker, R. J. (1999). *Multilevel analysis*. Sage Publications.

Snijders, T. A. B., & Bosker, R. J. (2011). *Multilevel analysis: An introduction to basic and advanced multilevel modeling*. Sage Publications.

Tajfel, H., & Turner, J. C. (1986). The social identity theory of group conflict. In S. Worchel & W. G. Austin (Eds.), *Psychology of intergroup relations* (pp. 7–24). Nelson-Hall.

Tetlock, P. E., Kristel, O. V., Elson, S. B., Green, M. C., & Lerner, J. S. (2000). The psychology of the unthinkable: Taboo trade-offs, forbidden base rates, and heretical counterfactuals. *Journal of Personality and Social Psychology*, *78*(5), 853–870. https://doi.org/10.1037/0022-3514.78.5.853

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. https://doi.org/10.18637/jss.v059.i05

Van Bavel, J. J., & Packer, D. J. (2021). *The power of us: Harnessing our shared identities to improve performance, increase cooperation, and promote social harmony*. Little, Brown Spark.

van Zomeren, M., d'Amore, C., Pauls, I. L., Shuman, E., & Leal, A. (2023). The intergroup value protection model: A theoretically integrative and dynamic approach to intergroup conflict escalation in democratic societies. *Personality and Social Psychology Review*. Advance online publication. https://doi.org/10.1177/10888683231192120

Wang, L. P., & Maxwell, S. E. (2015). On disaggregating between-person and within-person effects with longitudinal data using multilevel models. *Psychological Methods*, *20*(1), 63–83. https://doi.org/10.1037/met0000030

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

Wisneski, D. C., Hanson, B. E., & Morgan, G. S. (2020). The roles of disgust and harm perception in political attitude moralization. *Politics and the Life Sciences*, *39*(2), 215–227. https://doi.org/10.1017/pls.2020.22

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

Zaal, M. P., Laar, C. V., Ståhl, T., Ellemers, N., & Derks, B. (2011). By any means necessary: The effects of regulatory focus and moral conviction on hostile and benevolent forms of collective action. *British Journal of Social Psychology*, *50*(4), 670–689. https://doi.org/10.1111/j.2044-8309.2011.02069.x

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

Received May 17, 2023
Revision received December 1, 2023
Accepted December 14, 2023 ■