Article

Morality, Social Inclusion, and In-Group Superiority: The Differential Role of Individualizing and Binding Foundations in Perceptions of the Social Identity of In-Group and Out-Group Members

Personality and Social Psychology Bulletin
2025, Vol. 51(12) 2571–2585

DOI: 10.1177/01461672241262367
journals.sagepub.com/home/pspb


Adi Amit[1] and Eitan Venzhik[1]

Abstract

Research into social identity and morality judgments typically focuses on how the former influences the latter. We approach this theme from the opposite direction, establishing the influence of morality on perceptions of social identity. In three studies, conducted in two cultures, we show that in-group members acting immorally are excluded from the group. Extending this investigation to the overlooked study of out-group-on-out-group behavior, in Studies 2 and 3, we compare perceptions of social inclusion for in-group members following (im)moral behavior toward the in-group with perceptions of out-group members following (im)moral behavior toward the out-group. We show that people treat in-group and out-group members alike with respect to binding morals, which concern preservation and protection of the group, but not individualizing morals, which concern preservation and protection of individual rights. Finally, in Study 3, we confirm the underlying motivating mechanism of reasserting in-group superiority by affirming the positive distinctiveness of one's own group.

Keywords

social identity, moral foundations, in-group superiority, social exclusion, collective narcissism

Received September 1 2023; revised accepted May 28 2024

Social psychologists have recently addressed the role of morality in regulating intragroup behavior, pointing to groups as moral anchors (Ellemers, 2017; Ellemers et al., 2013; Ellemers & Van der Toorn, 2015). Emphasizing the importance of the rarely addressed link between social identity and morality, Spears (2021) notes that the few studies on this issue tend to address how group membership influences the way people determine what constitutes (im)moral behavior. In the current research, rather than looking at how group membership affects moral judgments, we focus on the opposite influence—*how different types of (im)moral behavior affect judgments of (another's) group membership*. For example, we ask to what extent Americans consider as American a person engaging in different immoral acts within the in-group (such as pushing an American girl off a swing or burning the American flag), especially in comparison to the parallel moral acts (saving an American girl from falling off a swing, or wrapping oneself in the American flag).

Moreover, little research investigates judgments about the conduct of *out-group members toward fellow out-group members*. People often exhibit bias in favor of in-group members while derogating out-group members when judging the behavior of others (Hewstone et al., 2002). However, to our knowledge, no scholars have asked how different types of behavior affect perceptions of another's group membership. Thus, we further compare perceptions about the group membership of out-group members engaging in similar (im)moral behaviors (e.g., how Israelis judge the group membership of Palestinian individuals described as pushing a Palestinian girl off a swing vs. protecting her). That is, we ask how out-group members' behavior affects (outside observers') perceptions of their membership in their own group.

Why do people's judgments of others' group identity matter? Such judgments have repercussions at various levels, from reinforcing stereotypes to attributions of guilt—for example, whether a person is viewed as accountable for his or her actions or whether the group should be viewed as responsible. Differences in such perceptions of in-group and

[1]The Open University of Israel, Ra'anana, Israel

Corresponding Author:
Adi Amit, The Open University of Israel, 1 University Road, Ra'anana 4353701, Israel.
Email: adiam@openu.ac.il

out-group members indicate a double standard that could help explain intolerance toward immigrants and even reduced acceptance of dual nationals or second-generation group members. Moreover, social exclusion has been proposed as a mechanism by which groups restore threatened group positivity following in-group deviance (e.g., Ditrich & Sassenberg, 2016; Eidelman et al., 2006; Jetten & Hornsey, 2014; Pinto et al., 2010). We suggest such processes could serve as a motivational mechanism also with regard to out-group deviance.

We suggest that individuals rely on morality as a prism through which they perceive and interpret social reality. We draw differential predictions regarding the perceived social identity of an (im)moral actor based on the nature of the (im)moral act and the actor's in-group/out-group membership. Relying on Moral Foundations Theory (MFT; Graham et al., 2013; Haidt & Joseph, 2004), we distinguish between adherence to (or violation of) moral principles that affect intra-group relations (binding principles) and those that relate to individual rights (individualizing principles). We show that people treat in-group and out-group members alike with respect to binding, but not individualizing, morals in ways that reassert in-group superiority.

Social Identity and Morality

Kesebir and Haidt (2010, p. 800) define morality as "interlocking sets of values, virtues, norms, practices, identities, institutions, technologies, and evolved psychological mechanisms that work together to suppress or regulate selfishness and make social life possible." This definition emphasizes the role of morality in regulating human conduct by defining what constitutes wrong versus right behavior. Thus, it relates to morality as a regulatory mechanism that maintains social order in groups and communities (Haidt, 2008; Janoff-Bulman & Carnes, 2013).

Theorists of morality tend to adopt an absolutist deontic sense of morality, overlooking the influence of group identity (Spears, 2021). For example, adopting Kantian ethics, Skitka and colleagues (2005) focus on moral conviction as transcending the boundaries of persons and cultures. Similarly, MFT explains how moral principles can have universal foundations while still varying significantly across cultures and societies (Graham et al., 2013; Haidt & Joseph, 2004). Such theories focus on the moral acts themselves and do not consider the social identity of those concerned—either the actors who defy moral convictions or principles or those who suffer from immoral acts.

However, the practice of excluding identity markers from moral scenarios fails to capture the social core of morality. It is not just that, as Hester and Gray (2020) note, to the degree that participants hail from the WEIRD (Western, Educated, Industrialized, Rich, and Democratic) group of nations, they are also likely to assume the characters in such scenarios fit a certain archetype, namely white, middle-aged men. It is that social identity itself plays a crucial role in how people view (im)moral behavior, with respect to the identity of the transgressor, the victim, or both. For example, research on norm violations focuses on how deviations from group norms affect other members of the group or the group as a whole (e.g., Livingstone et al., 2011; Van Kleef et al., 2011). Likewise, studies of intergroup transgressions show that harm inflicted on the in-group by out-group leaders is judged more severely than similar harm inflicted on an out-group by in-group leaders (e.g., Abrams et al., 2013) and that individuals favor harsher punishment for dissent when it is perceived as reflecting disloyalty—that is, motivated by a wish to help an out-group at the expense of the in-group (Roccas et al., 2022). Research comparing judgments toward in-group and out-group members suggests that deviant in-group members are judged more negatively than comparable deviant out-group members (e.g., Marques et al., 1988, 2001; Travaglino et al., 2014), punished more harshly (Mendoza et al., 2014), and even ostracized (Castano et al., 2002; Eidelman et al., 2006). Similarly, Bettache and colleagues (2019) show that higher identifiers judge in-group members who violate binding (though not individualizing) moral principles more harshly than out-group members committing the same offenses.

Importantly, most research mentioned earlier tends either to focus on the social identity (in-group/out-group) of the transgressor but not the target of the behavior or to consider the social identity of both parties with behavior directed at the other group. Such research has done much to expose different ways that people adjust their moral lens to neutralize threats to both their moral self-image and their social identity (e.g., Leidner & Castano, 2012). But few of these studies address judgments toward out-group members whose immoral behavior targets *their own* group. And those studies that do investigate moral judgments toward out-group-on-out-group behavior focus almost exclusively on judgments of fairness. For example, members of the in-group punished in-group members who made unfair allocations to fellow in-group members more harshly than out-group members who made unfair allocations to their fellow out-group members (Bernhard et al., 2006), and participants were more likely to recall unfair behavior toward fellow group members (e.g., via allocations in a trust game or descriptions of cheating) when the violator was a member of the in-group rather than the out-group (Hechler et al., 2016). Importantly, these findings also show that individuals are not indifferent to moral and immoral behavior by out-group members: Unfair allocations by out-group members are not left unpunished (Bernhard et al., 2006), and moral and immoral actions by out-group members are remembered better than neutral actions by both in-group and out-group members (Hechler et al., 2016). Note, though, that fairness is but one of several foundations of morality (Graham et al., 2013). Other types may be viewed differently.

MFT: Different Types of Morality

MFT is a prominent theory of morality that seeks to explain how moral principles can have universal foundations while still varying significantly across cultures and societies (Graham et al., 2013; Haidt & Joseph, 2004). MFT identifies five foundations of intuitive ethical principles (care, fairness, loyalty, authority, and purity), which are then subject to social and cultural adjustments that shape the unique moral code of each culture. A sixth foundation, liberty, was later added to the theory (Haidt, 2012; Iyer et al., 2012). Past research has shown that the foundations are stable across cultures (Doğruyol et al., 2019; Haidt & Graham, 2007) while also operating differently in different groups (e.g., Graham et al., 2009). Winget and Tindale (2020) suggest that MFT is particularly suited to study the links between morality and social identity because it (a) defines morality broadly, extending beyond traditional justice and rights; (b) is stable across groups, as it focuses on moral systems rather than the content of a given moral principle; and (c) highlights the importance of morality in maintaining social life.

According to Haidt and Graham (2007), moral foundations can be classified as *individualizing* or *binding. Individualizing* morals focus on individual rights. They include the ability to feel (and dislike) pain in others (the care/harm foundation) and ideas of justice, rights, and autonomy (the fairness/cheating foundation). *Binding* morals focus on the group. They relate to devotion to the group (loyalty/betrayal), deference to legitimate authority and respect for tradition (authority/subversion), and pursuit of an elevated (non-desecrated) life (purity/degradation).

Binding morals are related more than individualizing morals to group-oriented structures and belief systems, including right-wing authoritarianism (Federico et al., 2016; Kugler et al., 2014), conservatism (Koleva et al., 2012), cooperation (Clark et al., 2017), self-control (Mooijman et al., 2018), and need for cognitive closure (Federico et al., 2016). The role of the group in people's judgments about morality is striking when group identities justify overriding individualizing principles. This allows people to limit cooperative behavior to the in-group (Yamagishi & Kiyonari, 2000), or even to rationalize harming out-group members if it benefits the in-group (Smith et al., 2014).

Conceptualizations used in the study of norms can further highlight important differences between binding and individualizing foundations. Marques and colleagues (2001) differentiate between generic norms that apply to individuals of all groups, and particular norms that apply only to members of a certain group. For example, even non-Muslims should remove their shoes when entering a mosque, but only Muslims are expected to follow the Koran. The universal aspect of moral foundations suggests that all moral principles (individualizing and binding) should be viewed as generic. With respect to the preservation and protection of individual rights, individualizing foundations elicit similar behavioral expectations for everyone, regardless of group affiliation. Likewise, when speaking of the preservation and protection of the group, some expectations derived from binding foundations (like respecting the sanctity of religious spaces) constitute universal principles applicable to everyone. But other binding foundations, while still generic, also take on aspects of particularity. People are expected to be loyal, express appropriate forms of submission, and adhere to the sanctity rules of *their own group*, which differ between groups; failure to do so—for example, for Muslims, failure to adhere to the Koran—is only a moral violation for members of that group. Thus, even though the conceptualization of norms as generic or particular is differentially relevant to individualizing and binding moralities, the two distinctions are not equivalent.

Morality as a Mechanism for Affirming In-Group Superiority

We follow Ellemers and colleagues (2013) and Ellemers and van den Bos (2012) in emphasizing the intergroup regulatory function of morality. According to the Behavioral Regulation Model, morality is recognized as an important virtue of groups and a key source of group pride and identification (Abele et al., 2021). To maintain a positive self-image through identification with positively valued social groups, individuals are attracted to groups perceived as moral, such that morality, more than competence, determines whether individuals are attracted to particular organizations and work teams (van Prooijen & Ellemers, 2015; van Prooijen et al., 2018). Moreover, to maintain group positivity, members who adhere to moral standards are viewed as "good," whereas those who do not adhere to such standards risk social exclusion (Ellemers et al., 2013). Hence, beyond competence, newcomers are accepted or rejected in organizations based on their perceived moral attributes (Pagliaro et al., 2013; van der Lee et al., 2017). Thus, group members actively seek to bolster a sense of in-group superiority by embracing members who act morally and excluding those who do not.

Research on deviance supports the notion that reactions to (im)moral behavior may serve to restore the positivity of one's own group. In a review of the literature on deviance as a violation of group norms (not necessarily moral), Jetten and Hornsey (2014) suggest that restoring group positivity is one of five motives that underlie the rejection of deviance. Group members may respond to deviant behavior by either excluding the deviant from the group or leaving the group themselves (Ditrich & Sassenberg, 2016). A meta-analysis showed that people who deviate from group norms are more likely than not to be rejected by the group (Richard et al., 2003).

Studies have found exclusion of group members for deviation from a range of norms. These include withholding information from teammates or distributing resources

unfairly (Ditrich & Sassenberg, 2016); expressing non-normative opinions about ideological issues (Eidelman et al., 2006); or displaying behaviors that undermine the stereotype of the group (Castano et al., 2002). However, as research on social norms currently lacks a taxonomy of norm types, there is no basis for integrating insights derived from studies on one norm to other norms (van Kleef et al., 2019).

Within the realm of morality, by contrast, reliance on MFT allows us to concurrently test what happens when people violate different moral principles. We suggest that, like in-group members who deviate from group norms, in-group members who violate any moral foundation challenge the positivity of their group and so are more likely to be excluded compared with those who adhere to the moral foundations. We hence hypothesize that:

> **Hypothesis 1 (H1):** For all moral foundations, in-group members who behave in ways consistent with moral conduct will be perceived as included in the group more than those who behave in ways inconsistent with moral conduct.

Perceptions Toward Out-Group Members

As noted previously, only a few studies examine the perceived morality of acts committed by out-group members, and those tend to focus on differential perceptions and judgments of the same act committed by in-group versus out-group members. Such studies have shown that people's judgments are biased in ways that exemplify the role of morality as a psychological tool for maintaining social order in groups (Haidt, 2008). Morality, more than competence or sociability, is an important determinant of the distinction between the in-group and out-groups, and positive evaluations of the former (Leach et al., 2007). Hence, one should expect the perceived social inclusion of (im)moral actors to differ for in-group and out-group members.

Individualizing foundations are interpersonal in nature, with the locus of moral concern being other individuals, whereas binding foundations are collective in nature, with the locus of moral concern being the group (Janoff-Bulman & Carnes, 2013). As such, group members who violate binding foundations (loyalty, authority, purity) explicitly signal that they do not care about being considered part of their group; they may even be seen as deliberately excluding themselves from the group. Thus, others should assess such violators as being socially excluded from their group. This applies to in-group and out-group members alike. Indeed, research has shown that people expect even enemy out-group members to display loyalty to their group (Amit et al., 2024).

In contrast, violations of individualizing foundations, which concern interpersonal interactions, should be judged differently for in-group and out-group members. As discussed above, people are motivated to perceive their in-group as more moral than the out-group—an important means of maintaining a positive group self-image. Interestingly, while people may prefer to interact with moral members of out-groups as well as the in-group, their motivations for doing so differ: interacting with moral (rather than immoral) in-group members is motivated by the desire to maintain a positive group image, whereas interacting with moral (rather than immoral) out-group members is motivated by a desire to maintain group safety (Brambilla et al., 2013). Hence, in-group members who harm or are unfair to fellow group members (i.e., who violate individualizing foundations) should be seen as worthy of exclusion from the group—a necessary means to maintain the moral image of the in-group, and by extension a moral self-image. By contrast, out-group members are expected to be morally inferior; learning of moral violations by out-group members is self-fulfilling (Sacchi et al., 2021). It follows that out-group members who violate individualizing foundations should be perceived as included in the out-group just as much as (or even more than) out-group members who respect individualizing foundations.

> **Hypothesis 1 (H2):** We expect a three-way interaction such that the effect of moral (vs. immoral) behavior on perceived social inclusion (the degree to which the actor is perceived as a member of his or her group) will depend on the social identity of the actor (in-group vs. out-group) and the type of moral foundation concerned (individualizing vs. binding). We expect that the difference between group inclusion perceptions toward moral versus immoral in-group members will be similar for individualizing and binding foundations, but the difference for out-group members will be smaller for individualizing compared to binding foundations.

Individual Differences in Investment in In-Group Superiority

Individuals differ in the extent of their concern with in-group superiority. Golec de Zavala and colleagues (2009) used the term "collective narcissism" for emotional investment in unrealistic beliefs about the greatness of one's in-group. Collective narcissism has been studied mainly in regard to the national group, with scholars suggesting that it might be the basis of nationalism (Federico et al., 2023). Reflecting a need to affirm and maintain a sense of in-group superiority, high levels of collective narcissism predict a readiness to perceive out-groups as threatening, an unwillingness to forgive out-groups (Golec de Zavala et al., 2009), and expressions of contempt toward out-groups (Golec de Zavala et al., 2013).

Importantly, collective narcissism enhances the effect of social identity on moral judgments, such that differential judgments toward deviant in-group versus out-group members are especially strong among those high in collective narcissism (Bocian et al., 2021). Following this finding, and in

light of our rationale that differential perceptions of group inclusion are motivated by the need to maintain in-group superiority, we suggest that individual differences in that motivation (represented as collective narcissism) will affect perceptions of group inclusion concerning individualizing foundations, such that:

> **Hypothesis 3 (H3):** Focusing on individualizing foundations, we expect the difference between group inclusion perceptions toward moral versus immoral in-group members to be larger than the same difference for out-group members. We further expect that this interactive effect will increase as the perceiver's emotional investment in in-group superiority rises.

The Current Research

We present three studies examining the influence of (im)moral behavior on the perceived social identity (group inclusion) of in-group and, more importantly, out-group members. As this is among the first studies to examine how out-group-on-out-group behavior is perceived by in-group members, we operationalized in-group and out-group membership based on national groups, and in particular conflictual (Study 2) or competing (Study 3) national groups, where in-group superiority is expected to be an especially strong motivator. We operationalized social inclusion using measures of prototypicality or group–person overlap.

In Study 1, we examine the influence of (im)moral conduct on the inclusion or exclusion of in-group members, testing H1. In Studies 2 and 3, we extend our study to judgments about the conduct of out-group members toward fellow out-group members, looking at how adherence to (or violation of) different moral principles affects perceptions of actors as members of their (out)group. We rely on MFT to distinguish between (im)moral actions capturing individualizing moral foundations and those capturing binding foundations, testing for the differential predictions in H2. In Study 3, we further test H3, positing emotional investment in in-group superiority as the mechanism that underlies biased perceptions.

Transparency and Openness

For all studies, we report how we determined our sample size, all data exclusions (if any), all manipulations, and all measures. All data, analysis codes, and research materials are available at https://osf.io/8v97q/?view_only=20ac975a4452414689112dcbaa895643. Data were analyzed using SPSS version 28. In all studies, participation was anonymous and voluntary.

Study 1

We first focus on in-group members. We test whether, for all moral foundations, in-group members who behave morally will be viewed as members of the group more than those who behave immorally (H1). We further test whether the difference is due to social exclusion of the immoral actor or to inclusion of the moral one, by comparing the moral and immoral conditions to a morally neutral condition. To operationalize group membership we focus on the national level, looking at how Americans perceive other Americans described as performing (im)moral acts.

Method

Participants and Procedure. Participants were 301 American adults (mean age = 40.2, standard deviation [*SD*] = 13.197; 49.8% female) recruited in March 2022[1] using the Prolific Academic online survey service. Sensitivity analysis assuming a correlation of .2 between repeated measures (as reported in Hadarics & Kende, 2018) showed that a sample of 300 would allow for the detection of effect size as small as $f = .10$ with 80% power (see the Supplementary Materials, SM1.1). All participants answered an attention item correctly and were included in the analysis.

Participants were randomly allocated to one of the three between-subjects experimental conditions, in which they were presented with vignettes describing moral behavior, immoral behavior, or morally neutral behavior. All characters in the vignettes were presented as belonging to the participants' in-group (Americans). All participants were presented with the full set of vignettes covering all five moral foundations.

Measures and Materials. Participants were presented with 15 vignettes, three for each of the five original, commonly studied moral foundations. The vignettes in the immoral condition were inspired by previous studies. The vignettes in the moral condition were created to mirror the immoral ones. In addition, we created neutral vignettes depicting similar contexts without moral or immoral behavior. To reflect group identity, we provided typical American names and cities for the person performing the action and the person(s) affected by it. This method has been used in past studies focusing on other contexts (e.g., Ashburn-Nardo et al., 2001; Coman et al., 2014). In all three conditions, adequate internal consistencies were obtained for the aggregated individualizing and binding indices and for each of the five moral foundations (.765–.954). The sources and adaptations of the vignettes, with internal consistencies, are presented in SM1.2.

We measured group inclusion perceptions through the notion of prototypicality. Participants were asked to report on a five-point scale to what extent they thought the person depicted in each vignette was a true (i.e., prototypical) American, from 1 (*not at all American*) to 5 (*very much American*).

Results and Discussion

We used a mixed analysis of variance (ANOVA) with morality (moral, immoral, and morally neutral) as a

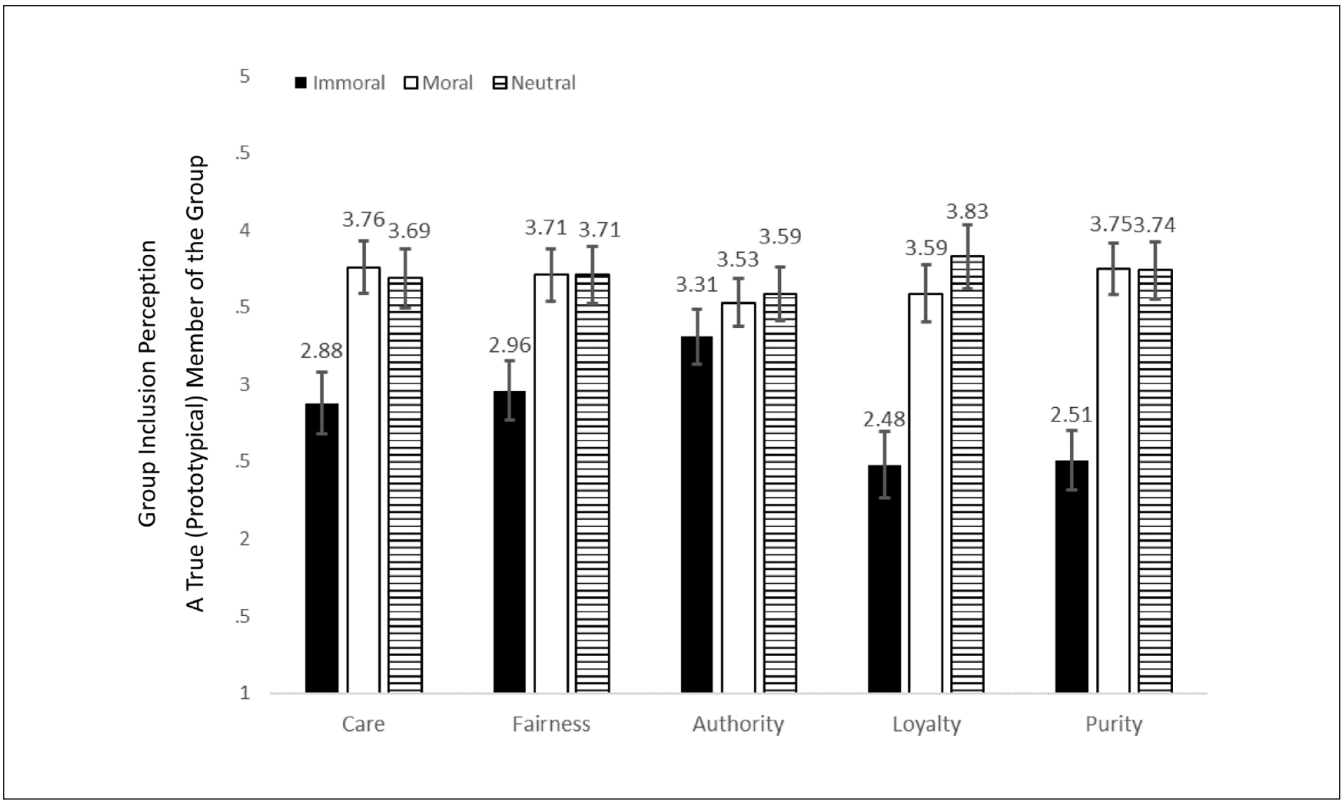

Figure 1. Means and Confidence Intervals of Social Inclusion Across All conditions, for Each of the Two Individualizing and Three Binding Moral Foundations (Study 1).

between-subjects factor and moral foundation (care, fairness, loyalty, authority, and purity) as a within-subjects factor. Means and confidence intervals are presented in Figure 1. To correct for unequal cell sizes, we report estimated marginal means. A non-hypothesized main effect for moral foundation emerged ($F(1, 298) = 7.909$, $p = .005$, using lower-bound adjustment due to violation of the sphericity assumption, $\chi^2(9) = 262.122$, $p < .001$). We also found a non-hypothesized interactive effect ($F(2, 298) = 23.184$, $p < .001$, $\eta_p^2 = .135$). More importantly, we found a significant main effect for morality ($M_{\text{moral}} = 3.669, M_{\text{immoral}} = 2.826, M_{\text{morally-neutral}} = 3.712$, $F(2, 298) = 33.44$, $p < .001$, $\eta_p^2 = .183$). Pairwise comparisons support our hypothesis that actors in the moral vignettes were perceived as members of the group more than those in the immoral condition ($M_{\text{Diff}} = 0.843$, $p < .001$, 95% confidence interval [CI] = [0.612, 1.074]). This was found for each of the five moral foundations (see SM1.3).

The difference between the moral and morally neutral vignettes was not significant ($M_{\text{Diff}} = -0.043$, $p = .706$, 95% CI = [−0.269, 0.183]). This could indicate either that it is the exclusion of actors behaving immorally rather than the inclusion of actors behaving morally that drives the above differences, or that moral behavior does not deviate from neutrality in the same way that immoral behavior does. Note, however, that the group identity of the actors in all conditions was established through the use of typical names and cities. Therefore, this lack of difference may also simply suggest that the names and cities are indeed indicative of the group identity. The actors in the morally neutral condition were perceived unequivocally as members of the group, to the extent that the lack of difference from the moral condition may be due to a ceiling effect.

To further compare the moral and immoral conditions, we combined the two individualizing foundations and the three binding foundations to create two indices (see Figure 2). To test for a potential differential effect of the type of moral foundation, we used a mixed ANOVA with morality (moral vs. immoral) as a between-subjects factor and foundation (binding or individualizing) as a within-subjects factor. To correct for unequal cell sizes, we report estimated marginal means. Supporting our hypothesis, we found a significant main effect for morality ($F_{(1,204)} = 48.531; p < .001, \eta_p^2 = .192$), such that in-group members who comply with moral principles were viewed as members of the group ($M = 3.680$) more than those who violate such principles ($M = 2.841$). We also found an additional significant effect of moral foundation type. Individualizing foundations were associated with higher social inclusion scores ($M = 3.327$) compared to the binding foundations ($M = 3.194$; $F(1,204) = 11.705; p < .001, \eta_p^2 = .054$).

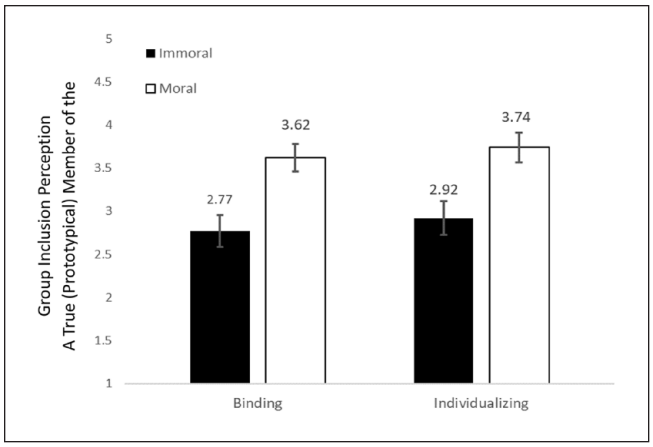

Figure 2. Means and Confidence Intervals of Social Inclusion for Moral Versus Immoral Behavior Reflecting Individualizing and Binding Moral Foundations (Study 1).

However, this result may be due to the large sample size for testing within-subject factors. No other significant effects were found. Table 1 presents this analysis in a summary of equivalent analyses for the three studies presented in this article.

Study 2

We next extend our examination of how (im)moral behavior influences perceived group identity to include out-group-on-out-group behavior, as compared with in-group-on-in-group behavior. Again, we focus on national groups. This time, the in-group members are Israelis and the out-group members are Palestinians. We test for differential expectations for behavior that violates individualizing versus binding foundations when out-group members are considered.

Method

Participants and Procedure. Participants were 471 Israeli adults (mean age = 45.244, *SD* = 16.718; 65.6% female) recruited in December 2021[2] using a local online survey service. We aimed to collect roughly 500 participants. Of the 506 participants recruited by the survey service, we eliminated 35 who failed to answer substantial parts of the survey (answering less than 3% of the questions). The remaining 471 participants correctly answered an attention item and were included in the final sample. Sensitivity analysis indicates that this sample size allows for the detection of an interactive effect size as small as $f = .10$ with 80% power (see SM2.1).

Participants were randomly allocated to one of four experimental conditions, in which they were presented with vignettes describing moral or immoral acts performed by and affecting individuals from the participants' in-group (Israelis) or out-group (Palestinians).

Measures and Materials. We used the moral and immoral vignettes and prototypical measures developed in Study 1, changing the names and cities according to the experimental condition (in-group or out-group). For the Israelis, we chose typical Jewish Israeli names and majority-Jewish Israeli cities, while for the Palestinians, we chose typical Palestinian names and cities located in the West Bank. In addition, we increased the number of items (see SM1.2).

Participants were asked to report on a five-point scale to what extent they thought the person depicted in each vignette was a true member of their group (Israeli in the in-group conditions and Palestinian in the out-group conditions). The scale ranged from 1 (*not at all Israeli/Palestinian*) to 5 (*very much Israeli/Palestinian*). Internal consistencies are presented in SM2.2.

Results and Discussion

Means and confidence intervals are presented in Figure 3. Focusing on the in-group conditions to test H1, the results fully replicate the findings of Study 1 (see Table 1 and SM2.3).

Focusing on the novel element of the current study (H2), we used a three-way mixed ANOVA with group identity (in-group vs. out-group) and morality (moral vs. immoral behavior) as between-subjects factors, and foundation type (binding or individualizing) as a within-subjects factor. As expected, we found a significant three-way interaction ($F_{(1,462)} = 7.377; p = .007, \eta_p^2 = .016$), confirming that how people perceive the group membership of moral versus immoral actors depends on the type of moral principle violated, as well as the actor's group identity.

Loyalty is most closely related to the specific content of group membership. Almost by definition, loyalty expresses a wish to be included in the group, and disloyalty expresses a wish to be excluded from the group. One may wonder whether the difference in social inclusion ratings between the moral versus immoral conditions for the individualizing but not binding foundations is driven specifically by loyalty. A series of six analyses, each comparing one of the individualizing foundations with one of the binding foundations, reduces this concern (see SM2.4).

Study 3

In Study 2, we confirmed that group inclusion expectations differ based both on whether the behavior violates individualizing versus binding moral foundations, and whether in-group or out-group members are considered. We next sought to establish the role of emotional investment in in-group superiority as the mechanism underlying the differential expectations toward in-group and out-group members who violate individualizing foundations. To further validate the generalizability of our findings, we use a different measure of group inclusion perceptions and extend our investigation

Table 1. Summary of Perceptions of Group Inclusion From Morality (Moral vs. Immoral) and Foundation (Binding vs. Individualizing) in Studies 1–3, and Comparisons Between the In-Group and Out-Group in Studies 2 and 3.

Topic	Effect	Study	*F*	*p*	η_p^2
Perceptions about In-group members	**H1: Main effect of morality (moral vs. immoral behavior)**	**1*** **	**48.531**	**<.001**	**.192**
		2* **	**161.211**	**<.001**	**.398**
		3* **	**107.736**	**<.001**	**.356**
	Main effect of foundation	1***	11.705	<.001	.054
		2*	6.197	.013	.025
		3***	133.459	<.001	.406
	Morality × Foundation	1	0.235	.0629	.001
		2***	57.772	<.001	.191
		3***	98.674	<.001	.336
Comparing perceptions about In-group and Out-group members	Main effect of morality	2***	146.927	<.001	.241
		3***	146.573	<.001	.278
	Main effect of foundation	2	2.473	.117	.005
		3***	142.416	<.001	.272
	Main effect of group	2	.021	.885	0
		3	1.123	.290	.003
	Morality x Foundation	2***	154.282	<.001	.250
		3***	199.587	<.001	.344
	Group x Morality	2***	19.403	<.001	.040
		3	.606	.437	.002
	Group x Foundation	2	2.579	.109	.006
		3	1.438	.231	.004
	H2: Three-way interaction	**2** **	**7.377**	**.007**	**.016**
		3 **	**8.038**	**.005**	**.021**

Note. Hypothesized Effects in Bold.

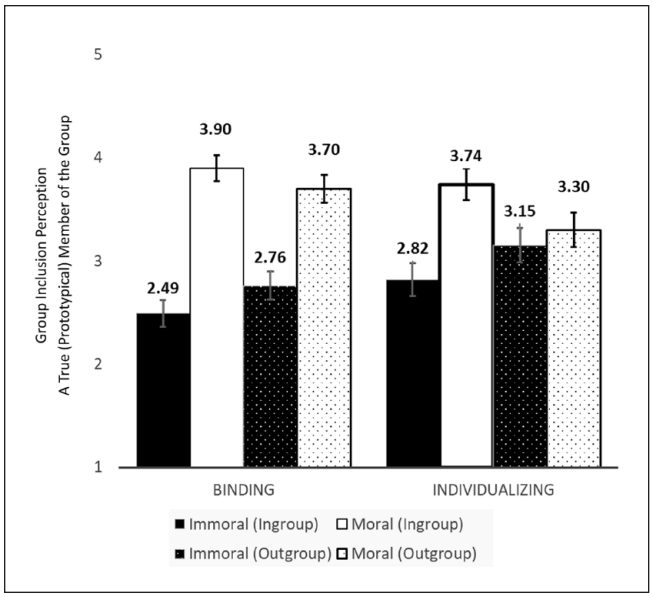

Figure 3. Means and Confidence Intervals of Social Inclusion for the Individualizing and Binding Foundations in All Experimental Conditions in Study 2.

to another national context—United States–China tensions. While this context is less acute than the Israeli–Palestinian conflict used in Study 2, the trade war between the two countries, as well as the spread of COVID-19, which originated in China, resulted in Americans holding unfavorable views of China (Devlin et al., 2020).

Method

Participants and Procedure. We aimed to collect roughly 600 participants to allow substantial power to detect the interactive effect with a new measure of group inclusion perceptions and to allow for exclusions (as preregistered at https://aspredicted.org/ZCG_75S).[3] Participants were 598 American adults, recruited in July 2023[4] using Prolific Academic, who correctly answered a general attention item[5] (mean age = 42.64, *SD* = 13.32; 40.9% female). We included training items to practice the new measure of group inclusion perception, in which a small circle representing a person could be moved in relation to a bigger circle representing a group (for more details see under Measures below and SM3.4). In the training items, participants were asked to move a small circle representing a person with a typical American name described as living in a typical American city toward or away from a larger circle representing the correct national group (American) or the incorrect national group (Chinese). The small circle could be positioned on a spectrum from fully enclosed in the larger circle (reflecting full social inclusion) to fully outside the larger circle (reflecting social exclusion). We removed from the analysis 213 participants who failed to

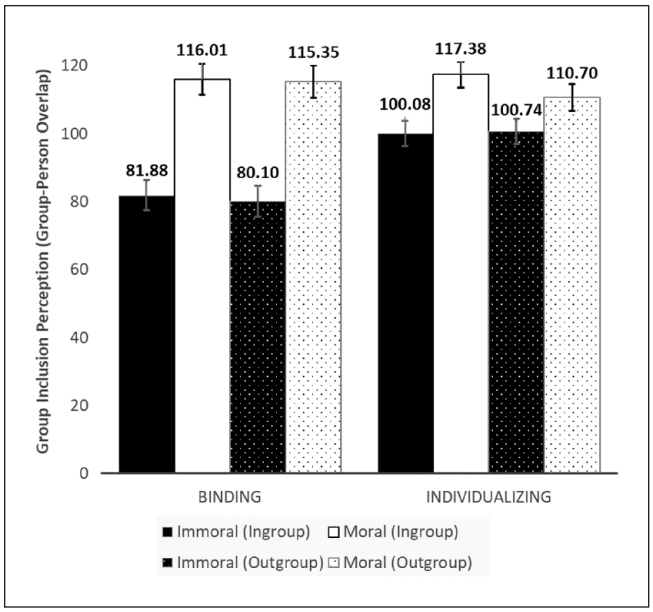

Figure 4. Means and Confidence Intervals of Group Inclusion Perceptions for the Individualizing and Binding Foundations in All Experimental Conditions in Study 3.

exclude the American person (small circle) from the Chinese group (big circle) or failed to include the American person (small circle) in the American group (big circle).[6] The remaining 385 participants were included in the final sample (mean age = 43.174, *SD* = 13.159; 45.2% female). Sensitivity analysis based on the actual sample size, 80% power, and $\alpha = .05$ indicated that the minimum detectable effect size is $f = .11$ (see SM3.1). We report here our results for the final sample. Analyses using the full sample are presented in SM3.3 (in those analyses, too, all hypotheses were fully confirmed).

To enhance the group manipulation we primed participants' national identity using a measure of national identification (Roccas et al., 2008), then measured investment in in-group superiority with the collective narcissism scale. Participants were then randomly allocated to one of four experimental conditions, based on vignettes describing moral or immoral acts performed by and affecting individuals from the participants' in-group (USA) or out-group (China).

Measures and Materials

Investment in In-group Superiority. We used the five-item scale of collective narcissism developed by Federico and colleagues (2023; e.g., "If the USA had a major say in the world, the world would be a much better place"). The measure yielded adequate internal reliability ($\alpha = .920$).[7]

Group Inclusion Perceptions. Participants were presented with the moral and immoral vignettes using typical American names and cities for the in-group conditions (as used in Study 1) and typical Chinese names and cities for the out-group conditions. We added three vignettes representing the newer, individualizing, foundation of liberty. To measure group inclusion, we used an adapted version of the Assessment of Person–Group Overlap tool (Schubert & Otten, 2002, originally used to measure inclusion of other in the self, Aron et al., 1992) numerically anchored at 0 for full exclusion and 125 for full inclusion. The slider scale and internal consistencies are presented in SM3.4.

Results and Discussion

Means and confidence intervals are presented in Figure 4. Testing H1 and H2, the results fully replicate the findings of Study 2 (see Table 1 and SM3.5).

Focusing on the novel contribution of the current study (H3), we used the PROCESS (Model 3) extension to SPSS (Hayes & Preacher, 2014) to test for a moderating effect of investment in in-group superiority on the difference between group inclusion perceptions toward in-group (coded as 0) and out-group (1) members who engage in moral (1) versus immoral (0) acts (looking only at the individualizing foundations). The expected three-way interaction was significant ($F_{(1,377)} = 4.142; p = .043, R^2_{change} = 0.009$), and is evident in the different slopes displayed in Figure 5 when comparing in-group and out-group members. Inclusion perceptions toward both in-group and out-group members who comply with individualizing moral foundations are not affected by attitudes toward in-group superiority. However, perceptions toward those who violate individualizing foundations differ, indicating greater sensitivity to in-group deviance: the greater participants' investment in in-group superiority, the more they exclude in-group members who violate individualizing foundations (conditional effect = −3.667, *SE* = 1.222, $t = -3.001$, $p = .003$, 95% CI = [−6.069, −1.264]). This pattern was reversed for out-group members, although the simple slopes fail to reach statistical significance (conditional effect = 1.361, *SE* = 1.171, $t = 1.162$, $p = .246$, 95% CI = [−0.942, 3.664]).

General Discussion

Scholars have rarely addressed the influence of different types of (im)moral behavior on judgments about group membership, let alone judgments based on the conduct of out-group members toward fellow out-group members. Thus, while Ellemers and Van der Toorn (2015) suggest that "universal moral guidelines are seen through the lens of group defining values" (p. 190), we suggest that group identity is seen (and attributed) through the lens of moral guidelines.

We show that immoral behavior by in-group members leads to social exclusion. This exclusion is evident for diverse types of morality and is similar in magnitude for the individualizing and binding moral foundations. These findings were consistent across three studies, with different populations

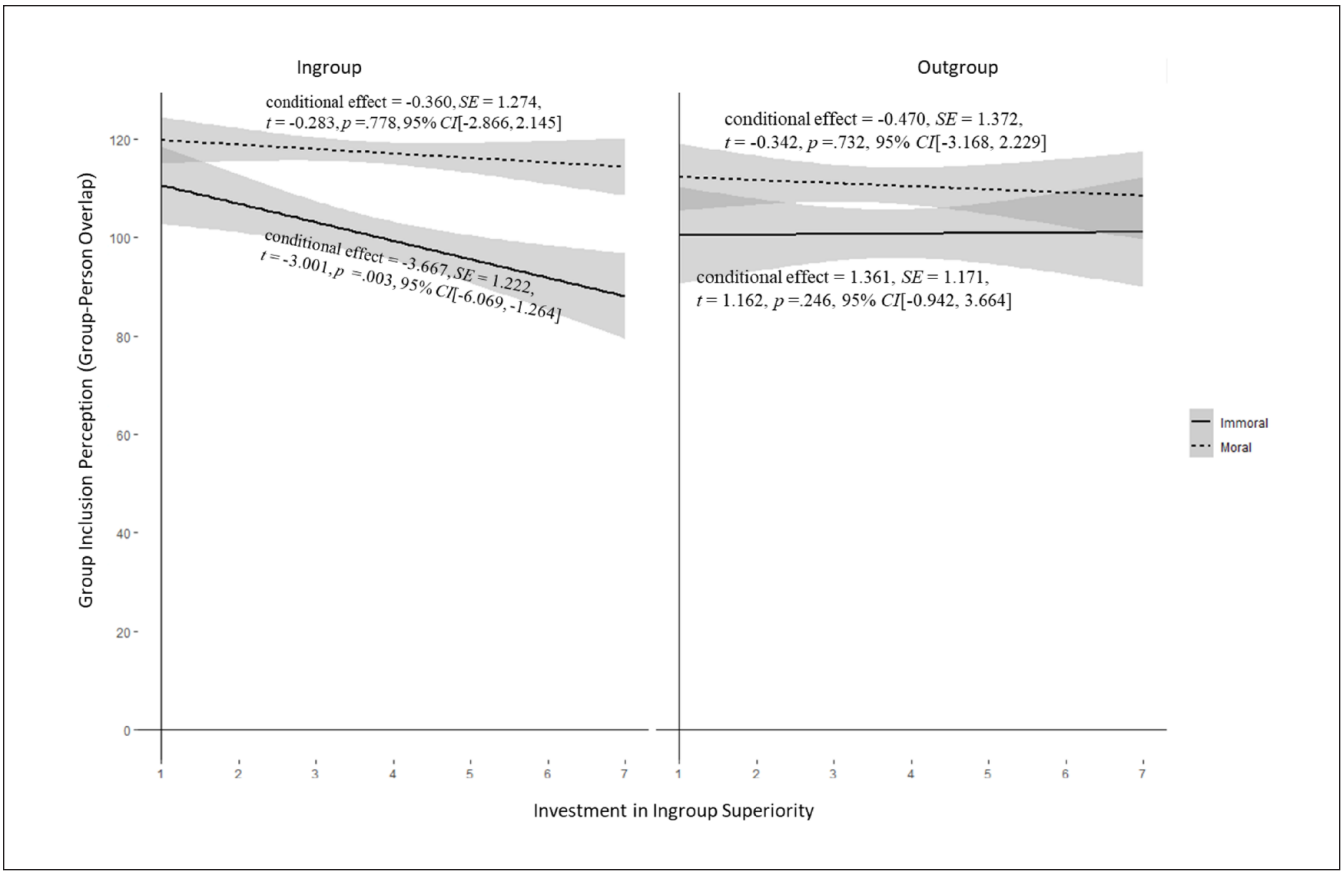

Figure 5 Effect of Investment in In-group Superiority on Group Inclusion Perceptions Toward Moral Versus Immoral In-group or Out-group Members in Study 3 (Looking Only at the Individualizing Foundations)
Note. Shaded areas represent the standard errors (*SEs*) around the regression lines.

from different national groups (Americans and Israelis). Extending the study to perceptions of out-groups, in Studies 2 and 3 we showed that perceptions of group inclusion toward out-group members depend on the type of morality. Out-group members who violate binding moral foundations were deemed less prototypical members of their group (in Studies 1 and 2) or were represented as overlapping less with their group (Study 3), compared with those who comply with binding foundations (similar to the situation for immoral in-group members). This was not the case for out-group members who violate individualizing moral foundations. In Study 3, we further confirm the role of in-group superiority in moderating perceptions toward out-group members who violate versus comply with individualizing foundations, compared with in-group members.

Social identity theorists argue that social categorization stimulates comparisons between the in-group and out-group, and that people are motivated to resolve these comparisons in ways that preserve (or enhance) their self-esteem by creating a representation of their group as superior to other groups. Our findings are in line with this rationale, and the underlying motivation was further confirmed in Study 3 via use of the collective narcissism scale. Determining the precise nature and components of in-group superiority is beyond the scope of this study, although it may be rooted in need for positive distinctiveness. Future research could follow dos Santos and Pereira (2021) to establish empirically the role of need for positive distinctiveness. Yet caution should be taken in teasing apart differentiation from distinctiveness (Jetten & Spears, 2003).

Our findings confirm that social exclusion is a mechanism for maintaining in-group superiority, and is therefore applied differently for in-group and out-group deviants. We thus extend past theorizing focusing on the role of social exclusion in restoring group positivity threatened by *in-group* deviants (see Jetten & Hornsey, 2014). The notion that in-group positivity can be restored by social exclusion underlies the black sheep effect (Marques et al., 2001), in which judgments toward deviant in-group members are harsher than judgments toward deviant out-group members. Interestingly, another way to bolster belief in the superiority of the in-group is to downplay in-group immorality—that is, to judge in-group deviance more leniently (moral hypocrisy, Valdesolo & DeSteno, 2007). Research is yet to resolve which of these two competing mechanisms is more likely to

be used and under what conditions. Perhaps moving from moral judgments to perceptions of group inclusion could help tease apart the two. Future studies could also test whether exclusion reduces the need to judge violators harshly, and vice versa.

We focused our study on the moral realm: reactions to behavior such as harming others, behaving unfairly, disobeying authority, or acting disloyally. One may wonder whether the biased reactions found are also evident in other domains that concern acceptable and unacceptable behavior. For example, resting one's feet on a chair in a cafeteria or dropping ashes on the floor are considered unacceptable norm violations (Van Kleef et al., 2011), yet they are not immoral acts. Even young children distinguish moral from conventional norm violations, judging the former more severely (Smetana et al., 2018). Interestingly, children also derive differential expectations for norm abidance, expecting moral norms to be applicable to everyone and conventional norms to be applicable only to members of their community (Karadağ & Soley, 2023). As perceptions of norm violation foster ostracizing behavior (Rudert et al., 2023), future research may seek to replicate our findings with conventional norm violations.

One may wonder whether individualizing and binding moralities map onto the distinctions suggested by Abrams and colleagues (2014, 2017) between generic norms that apply to individuals across groups and oppositional norms that specify different, and even contrasting, behaviors for out-group members. However, the distinction between generic and non-generic norms is not yet well-defined, with loyalty, for example, considered a strong oppositional norm that also involves features of generic norms (Travaglino et al., 2014). This is especially interesting as in-group favoritism in itself is regarded as a generic norm (Iacoviello & Spears, 2018) that calls for a parochial view of loyalty (Amit et al., 2024). Relating to generic and particular norms instead of the content-based distinction between individualizing and binding moralities, one may expect similar responses to actions by all individuals, regardless of their group, especially for individualizing foundations, as these are not group-related and are thus more generic in nature. However, our findings support our theorizing linking the motivation to maintain in-group superiority to differential perceptions regarding the group inclusion of in-group and out-group members who conform to versus deviate from individualizing moralities.

We reasoned that the differences we observed in perceptions of national identity are driven by the classic in-group–out-group distinction. We focused our investigation on membership in conflictual (Study 2) or competing (Study 3) national groups in which in-group superiority is expected to be an especially strong motivator. This may limit the generalizability of our conclusions. It is possible that differential exclusion of out-group members is only evident in intergroup contexts in which the self-image of the in-group is threatened, and may not manifest when the out-group is an ally. Future studies could investigate other nations, or other types of groups that may differ in their importance to members' social identity as well as other characteristics (e.g., sports fans, ethnic groups, and even artificial groups created in the lab). This can be done while mapping potential moderators, such as cultural distance or status.

We relied on the construct of collective narcissism to conceptualize and measure emotional investment in in-group superiority. While different from nationalism, collective narcissism too was developed and largely investigated with regard to national groups (Cichocka & Cislak, 2020; Golec de Zavala, 2018). However, the motivation to seek and maintain in-group superiority is not confined to the national context and is even evident in lab-formed groups (e.g., Ellemers et al., 1999). Interestingly, Cichocka (2016) suggested that group idealization can manifest in two distinct forms: narcissistic, defensive in-group positivity, wherein a belief in in-group greatness is contingent on external validation; and secure in-group positivity, which is independent of how the group appears in the eyes of others. In light of this distinction, it might be that collective narcissism should be seen as representing in-group superiority only in its defensive form, which is also more closely linked to sensitivity to group status. Indeed, collective narcissism increases when group status is challenged (Bagci et al., 2023; Cichocka & Cislak, 2020), suggesting that differences in perceived group inclusion may be larger in samples composed of disadvantaged groups.

In addition to confirming our hypotheses, we found some consistent non-hypothesized effects (see Table 1). Focusing on the in-group alone, in all three studies, we find that behavior that exemplifies an individualizing (vs. binding) category yields greater social inclusion. One possible explanation for this finding is that because binding moral foundations are so closely linked to the group, they yield high social inclusion ratings for moral behavior alongside extremely low social inclusion ratings for immoral behavior. When the two behavior conditions are considered together, the combination of these two effects then yields an artificially low social inclusion result. This explanation is in line with the significant two-way interaction found in Studies 2 and 3, which shows that the difference in group inclusion scores for moral versus immoral behaviors reflecting *binding* moral foundations is larger than the equivalent difference for *individualizing* moral foundations.

To test our predictions, we relied on a set of moral vignettes. This common practice in the study of morality is limited by the choice of specific vignettes. To reduce these limitations, we included several vignettes for each moral foundation, carefully selected based on prior research (see SM1.2). The moral vignettes were created to equate the immoral ones, and indeed the difference in group inclusion perceptions between the moral and immoral vignettes (H1) was consistent for each of the five moral foundations in

Study 1, and across populations in Studies 2 and 3. It is therefore unlikely to result from the specific choice of vignettes. Binding and individualizing moralities differ in content and may provoke reactions to differing degrees (as suggested by the consistent non-hypothesized effect of the moral foundation discussed above). Importantly, our hypothesis on differential perceptions toward in-group and out-group members (H2) was tested using identical vignettes except for the names and cities indicating group membership. Thus, the interactive effect with group identity cannot be ascribed to differences driven by the specific set of vignettes. The lack of difference between the moral and neutral conditions in Study 1 suggests that moral behavior is perceived as normative behavior. It is possible that the effect of the group would have been even more pronounced if our vignettes had described "over the top" (i.e., deviant) moral behavior, as found when comparing pro-norm and anti-norm deviants (Abrams et al., 2000).

Overall, the current research sheds light on an overlooked link between morality and social inclusion, extending the investigation to the study of out-group-on-out-group behavior. We focused our research on perceptions of group membership given an individual's (im)moral behavior. Our findings confirm that people treat in-group and out-group members alike when it comes to binding moral foundations, but not when it comes to individualizing moral foundations. We hope our study will spark more research on judgments of intra-out-group phenomena.

Acknowledgments

The authors thank Meira Ben-Gad and Yuval Karmel for their thoughtful comments and suggestions.

Author Contributions

CRediT: A.A. contributed to conceptualization, methodology, writing (original draft), supervision, and funding acquisition. E.V. contributed to methodology, formal analysis, and visualization.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research was supported by a grant from the Israel Science Foundation to the first author (604/20).

Ethical Approval

All studies were approved by the ethics committee of the psychology department at the Open University of Israel.

ORCID iD

Adi Amit https://orcid.org/0000-0002-6497-7457

Data Availability Statement

All research materials, data, and analysis code are available at https://osf.io/8v97q/?view_only=191bc17c2857491fa40af543f97ebebd.

Supplemental Material

Supplemental material is available online with this article.

Notes

1. Data were collected shortly after Russia's invasion of Ukraine, at a time when China blamed the United States for the war. https://en.wikipedia.org/wiki/China%E2%80%93United_States_relations
2. Data were collected during a fairly stable phase of the enduring Israeli–Palestinian conflict, 6 months after the 2021 crisis and 8 months before the 2022 clashes. https://en.wikipedia.org/wiki/Gaza%E2%80%93Israel_conflict#2007%E2%80%932022
3. In the preregistered analysis plan, we declared that we would test three hypotheses. They were all tested and confirmed. However, we realized that the second registered hypothesis (focusing solely on the out-group conditions) is redundant, as it is embedded in the third hypothesis (expecting a different pattern for in-group and out-group members). We therefore do not report it here (or in Study 2).
4. Data were collected during a period of talks aiming to deescalate the conflict between the two countries and stabilize their relations. https://en.wikipedia.org/wiki/China%E2%80%93United_States_relations.
5. We used the day-of-week attention item. Participation of those who failed to mark the requested day was terminated immediately.
6. We also used two similar items in which the small circle represented a Chinese person with a typical Chinese name described as living in a typical Chinese city. Many participants failed to exclude the Chinese person from the (wrong) American group or to include that person in the (right) Chinese group. Some participants used the messaging system in Prolific to share that the person could be Chinese-American, and so could have a Chinese name and origin but nonetheless be fully American. We thus deviated from the preregistered plan and did not remove participants who failed those items. Analyses using the preregistered sample are presented in SM3.2.
7. The measure was strongly correlated with the superiority mode of identification ($r(385) = .865$, $p < .001$) measured by four ($\alpha = .885$) of the 16 identification items developed by Roccas and colleagues (2008). Past research has noted that the two constructs are conceptually and empirically similar (e.g., Biddlestone et al., 2022; Schori-Eyal et al., 2015). Since the superiority mode of identification has previously been used to measure perceived in-group superiority (Vollhardt et al., 2021), the high correlation validates our decision to use the collective narcissism scale.

References

Abele, A. E., Ellemers, N., Fiske, S. T., Koch, A., & Yzerbyt, V. (2021). Navigating the social world: Toward an integrated framework for evaluating self, individuals, and groups. *Psychological Review*, *128*(2), 290–314.

Abrams, D., Marques, J. M., Bown, N., & Henson, M. (2000). Pro-norm and anti-norm deviance within and between groups. *Journal of Personality and Social Psychology*, *78*(5), 906–912.

Abrams, D., Palmer, S. B., Rutland, A., Cameron, L., & Van de Vyver, J. (2014). Evaluations of and reasoning about normative and deviant ingroup and outgroup members: Development of the black sheep effect. *Developmental Psychology*, *50*(1), 258.

Abrams, D., Palmer, S. B., Van de Vyver, J., Hayes, D., Delaney, K., Guarella, S., & Purewal, K. (2017). Adolescents' judgments of doubly deviant peers: Implications of intergroup and intragroup dynamics for disloyal and overweight group members. *Social Development*, *26*(2), 310–328.

Abrams, D., Randsley de Moura, G., & Travaglino, G. A. (2013). A double standard when group members behave badly: Transgression credit to ingroup leaders. *Journal of Personality and Social Psychology*, *105*(5), 799–815.

Amit, A., Ben David, Y., & Venzhik, E. (2024). Expectations of national loyalty in members of nations other than one's own: Universalistic and parochial-based perspectives [Unpublished manuscript].

Aron, A., Aron, E. N., & Smollan, D. (1992). Inclusion of other in the self scale and the structure of interpersonal closeness. *Journal of Personality and Social Psychology*, *63*(4), 596.

Ashburn-Nardo, L., Voils, C. I., & Monteith, M. J. (2001). Implicit associations as the seeds of intergroup bias: How easily do they take root? *Journal of Personality and Social Psychology*, *81*(5), 789–799.

Bagci, S. C., Stathi, S., & Golec de Zavala, A. (2023). Social identity threat across group status: Links to psychological well-being and intergroup bias through collective narcissism and ingroup satisfaction. *Cultural Diversity and Ethnic Minority Psychology*, *29*(2), 208.

Bernhard, H., Fischbacher, U., & Fehr, E. (2006). Parochial altruism in humans. *Nature*, *442*(7105), 912–915.

Bettache, K., Hamamura, T., Amrani Idrissi, J., Amenyogbo, R. G., & Chiu, C. Y. (2019). Monitoring moral virtue: When the moral transgressions of in-group members are judged more severely. *Journal of Cross-Cultural Psychology*, *50*(2), 268–284.

Biddlestone, M., Cichocka, A., Główczewski, M., & Cislak, A. (2022). Their own worst enemy? Collective narcissists are willing to conspire against their in-group. *British Journal of Psychology*, *113*(4), 894–916.

Bocian, K., Cichocka, A., & Wojciszke, B. (2021). Moral tribalism: Moral judgments of actions supporting ingroup interests depend on collective narcissism. *Journal of Experimental Social Psychology*, *93*, 104098.

Brambilla, M., Sacchi, S., Pagliaro, S., & Ellemers, N. (2013). Morality and intergroup relations: Threats to safety and group image predict the desire to interact with outgroup and ingroup members. *Journal of Experimental Social Psychology*, *49*(5), 811–821.

Castano, E., Paladino, M. P., Coull, A., & Yzerbyt, V. Y. (2002). Protecting the ingroup stereotype: Ingroup identification and the management of deviant ingroup members. *British Journal of Social Psychology*, *41*(3), 365–385.

Cichocka, A. (2016). Understanding defensive and secure in-group positivity: The role of collective narcissism. *European Review of Social Psychology*, *27*(1), 283–317.

Cichocka, A., & Cislak, A. (2020). Nationalism as collective narcissism. *Current Opinion in Behavioral Sciences*, *34*, 69–74.

Clark, C. B., Swails, J. A., Pontinen, H. M., Bowerman, S. E., Kriz, K. A., & Hendricks, P. S. (2017). A behavioral economic assessment of individualizing versus binding moral foundations. *Personality and Individual Differences*, *112*, 49–54.

Coman, A., Stone, C. B., Castano, E., & Hirst, W. (2014). Justifying atrocities: The effect of moral-disengagement strategies on socially shared retrieval-induced forgetting. *Psychological Science*, *25*(6), 1281–1285.

Devlin, K., Silver, L., & Huang, C. (2020, April 21). *U.S. views of China increasingly negative amid coronavirus outbreak*. Pew Research Center, Global Attitudes and Trends. https://www.pewresearch.org/global/2020/04/21/u-s-views-of-china-increasingly-negative-amid-coronavirus-outbreak/

Ditrich, L., & Sassenberg, K. (2016). It's either you or me! Impact of deviations on social exclusion and leaving. *Group Processes & Intergroup Relations*, *19*(5), 630–652.

Doğruyol, B., Alper, S., & Yilmaz, O. (2019). The five-factor model of the moral foundations theory is stable across WEIRD and non-WEIRD cultures. *Personality and Individual Differences*, *151*(1), 109547.

dos Santos, M. F., & Pereira, C. R. (2021). The social psychology of a selective national inferiority complex: Reconciling positive distinctiveness and system justification. *Journal of Experimental Social Psychology*, *95*, 104–118.

Eidelman, S., Silvia, P. J., & Biernat, M. (2006). Responding to deviance: Target exclusion and differential devaluation. *Personality and Social Psychology Bulletin*, *32*(9), 1153–1164.

Ellemers, N. (2017). *Morality and the regulation of social behavior: Groups as moral anchors*. Routledge.

Ellemers, N., Kortekaas, P., & Ouwerkerk, J. W. (1999). Self-categorisation, commitment to the group and group self-esteem as related but distinct aspects of social identity. *European Journal of Social Psychology*, *29*(2–3), 371–389.

Ellemers, N., Pagliaro, S., & Barreto, M. (2013). Morality and behavioural regulation in groups: A social identity approach. *European Review of Social Psychology*, *24*, 160–193.

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.

Ellemers, N., & Van der Toorn, J. (2015). Groups as moral anchors. *Current Opinion in Psychology*, *6*, 189–194.

Federico, C. M., Ekstrom, P., Reifen-Tagar, M., & Williams, A. L. (2016). Epistemic motivation and the structure of moral intuition: Dispositional need for closure as a predictor of individualizing and binding morality. *European Journal of Personality*, *30*(3), 227–239.

Federico, C. M., Golec de Zavala, A., & Bu, W. (2023). Collective narcissism as a basis for nationalism. *Political Psychology*, *44*(1), 177–196.

Golec de Zavala, A. G. (2018). Collective narcissism: Antecedents and consequences of exaggeration of the in-group image. In A. Hermann, A. Brunell, & J. Foster (Eds.), *The handbook of trait narcissism: Key advances, research methods, and controversies* (pp. 79–89). Springer.

Golec de Zavala, A. G., Cichocka, A., & Bilewicz, M. (2013). The paradox of in-group love: Differentiating collective narcissism advances understanding of the relationship between in-group and out-group attitudes. *Journal of Personality*, *81*(1), 16–28.

Golec de Zavala, A. G., Cichocka, A., Eidelson, R., & Jayawickreme, N. (2009). Collective narcissism and its social consequences. *Journal of Personality and Social Psychology*, *97*(6), 1074–1096.

Graham, J., Haidt, J., Koleva, S., Motyl, M., Iyer, R., Wojcik, S. P., & Ditto, P. H. (2013). Moral foundations theory: The pragmatic validity of moral pluralism. *Advances in Experimental Social Psychology*, *47*, 55–130.

Graham, J., Haidt, J., & Nosek, B. A. (2009). Liberals and conservatives rely on different sets of moral foundations. *Journal of Personality and Social Psychology*, *96*(5), 1029–1046.

Hadarics, M., & Kende, A. (2018). The dimensions of generalized prejudice within the dual-process model: The mediating role of moral foundations. *Current Psychology*, *37*(4), 731–739.

Haidt, J. (2008). Morality. *Perspectives on Psychological Science*, *3*(1), 65–72.

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

Haidt, J., & Graham, J. (2007). When morality opposes justice: Conservatives have moral intuitions that liberals may not recognize. *Social Justice Research*, *20*(1), 98–116.

Haidt, J., & Joseph, C. (2004). Intuitive ethics: How innately prepared intuitions generate culturally variable virtues. *Daedalus*, *133*(4), 55–66.

Hayes, A. F., & Preacher, K. J. (2014). Statistical mediation analysis with a multicategorical independent variable. *British Journal of Mathematical and Statistical Psychology*, *67*(3), 451–470.

Hechler, S., Neyer, F. J., & Kessler, T. (2016). The infamous among us: Enhanced reputational memory for uncooperative ingroup members. *Cognition*, *157*, 1–13.

Hester, N., & Gray, K. (2020). The moral psychology of raceless, genderless strangers. *Perspectives on Psychological Science*, *15*(2), 216–230.

Hewstone, M., Rubin, M., & Willis, H. (2002). Intergroup bias. *Annual Review of Psychology*, *53*(1), 575–604.

Iacoviello, V., & Spears, R. (2018). "I know you expect me to favor my ingroup": Reviving Tajfel's original hypothesis on the generic norm explanation of ingroup favoritism. *Journal of Experimental Social Psychology*, *76*, 88–99.

Iyer, R., Koleva, S., Graham, J., Ditto, P. H., & Haidt, J. (2012). Understanding libertarian morality: The psychological roots of an individualist ideology. *PLOS ONE*, *7*(8), Article e42366.

Janoff-Bulman, R., & Carnes, N. C. (2013). Surveying the moral landscape: Moral motives and group-based moralities. *Personality and Social Psychology Review*, *17*(3), 219–236.

Jetten, J., & Hornsey, M. J. (2014). Deviance and dissent in groups. *Annual Review of Psychology*, *65*, 461–485.

Jetten, J., & Spears, R. (2003). The divisive potential of differences and similarities: The role of intergroup distinctiveness in intergroup differentiation. *European Review of Social Psychology*, *14*(1), 203–241.

Karadağ, D., & Soley, G. (2023). Children intend to teach conventional but not moral norms selectively to ingroup members. *Developmental Psychology*, *59*(3), 567.

Kesebir, S., & Haidt, J. (2010). Morality. In S. Fiske, D. Gilbert, & G. Lindzey (Eds.), *Handbook of social psychology* (5th ed., pp. 797–832). Wiley.

Koleva, S. P., Graham, J., Iyer, R., Ditto, P. H., & Haidt, J. (2012). Tracing the threads: How five moral concerns (especially Purity) help explain culture war attitudes. *Journal of Research in Personality*, *46*(2), 184–194.

Kugler, M., Jost, J. T., & Noorbaloochi, S. (2014). Another look at moral foundations theory: Do authoritarianism and social dominance orientation explain liberal–conservative differences in "moral" intuitions? *Social Justice Research*, *27*(4), 413–431.

Leach, C. W., Ellemers, N., & Barreto, M. (2007). Group virtue: The importance of morality (vs. competence and sociability) in the positive evaluation of in-groups. *Journal of Personality and Social Psychology*, *93*, 234–249.

Leidner, B., & Castano, E. (2012). Morality shifting in the context of intergroup violence. *European Journal of Social Psychology*, *42*, 82–91.

Livingstone, A. G., Haslam, S. A., Postmes, T. O. M., & Jetten, J. (2011). "We are, therefore we should": Evidence that in-group identification mediates the acquisition of in-group norms. *Journal of Applied Social Psychology*, *41*(8), 1857–1876.

Marques, J. M., Abrams, D., & Serôdio, R. G. (2001). Being better by being right: Subjective group dynamics and derogation of in-group deviants when generic norms are undermined. *Journal of Personality and Social Psychology*, *81*(3), 436–447.

Marques, J. M., Yzerbyt, V. Y., & Leyens, J. P. (1988). The "black sheep effect": Extremity of judgments towards ingroup members as a function of group identification. *European Journal of Social Psychology*, *18*(1), 1–16.

Mendoza, S. A., Lane, S. P., & Amodio, D. M. (2014). For members only: Ingroup punishment of fairness norm violations in the ultimatum game. *Social Psychological and Personality Science*, *5*(6), 662–670.

Mooijman, M., Meindl, P., Oyserman, D., Monterosso, J., Dehghani, M., Doris, J. M., & Graham, J. (2018). Resisting temptation for the good of the group: Binding moral values and the moralization of self-control. *Journal of Personality and Social Psychology*, *115*(3), 585–599.

Pagliaro, S., Brambilla, M., Sacchi, S., D'Angelo, M., & Ellemers, N. (2013). Initial impressions determine behaviours: Morality predicts the willingness to help newcomers. *Journal of Business Ethics*, *117*(1), 37–44.

Pinto, I. R., Marques, J. M., Levine, J. M., & Abrams, D. (2010). Membership status and subjective group dynamics: Who triggers the black sheep effect? *Journal of Personality and Social Psychology*, *99*(1), 107–119.

Richard, F. D., Bond, C. F., Jr., & Stokes-Zoota, J. J. (2003). One hundred years of social psychology quantitatively described. *Review of General Psychology*, *7*(4), 331–363.

Roccas, S., Amit, A., Oppenheim-Weller, S., Hazan, O., & Sagiv, L. (2022). Inclusive and exclusive beneficiary attributions: The role of social identity complexity in interpretations of and punishment for dissent. *Group Processes & Intergroup Relations*, *25*(6), 1653–1671.

Roccas, S., Sagiv, L., Schwartz, S., Halevy, N., & Eidelson, R. (2008). Toward a unifying model of identification with groups: Integrating theoretical perspectives. *Personality and Social Psychology Review*, *12*(3), 280–306.

Rudert, S. C., Möring, J. N., Kenntemich, C., & Büttner, C. M. (2023). When and why we ostracize others: Motivated social exclusion in group contexts. *Journal of Personality and Social Psychology*, *125*, 803–826.

Sacchi, S., Brambilla, M., & Graupmann, V. (2021). Basking in detected vice: Outgroup immorality enhances self-view. *Group Processes & Intergroup Relations*, *24*(3), 371–387.

Schori-Eyal, N., Tagar, M. R., Saguy, T., & Halperin, E. (2015). The benefits of group-based pride: Pride can motivate guilt

in intergroup conflicts among high glorifiers. *Journal of Experimental Social Psychology*, *61*, 79–83.

Schubert, T. W., & Otten, S. (2002). Overlap of self, ingroup, and outgroup: Pictorial measures of self-categorization. *Self and Identity*, *1*(4), 353–376.

Skitka, L. J., Bauman, C. W., & Sargis, E. G. (2005). Moral conviction: Another contributor to attitude strength or something more? *Journal of Personality and Social Psychology*, *88*(6), 895–917.

Smetana, J. G., Ball, C. L., Jambon, M., & Yoo, H. N. (2018). Are young children's preferences and evaluations of moral and conventional transgressors associated with domain distinctions in judgments? *Journal of Experimental Child Psychology*, *173*, 284–303.

Smith, I. H., Aquino, K., Koleva, S., & Graham, J. (2014). The moral ties that bind. . . even to out-groups: The interactive effect of moral identity and the binding moral foundations. *Psychological Science*, *25*(8), 1554–1562.

Spears, R. (2021). Social influence and group identity. *Annual Review of Psychology*, *72*, 367–390.

Travaglino, G. A., Abrams, D., de Moura, G. R., Marques, J. M., & Pinto, I. R. (2014). How groups react to disloyalty in the context of intergroup competition: Evaluations of group deserters and defectors. *Journal of Experimental Social Psychology*, *54*, 178–187.

Valdesolo, P., & DeSteno, D. (2007). Moral hypocrisy: Social groups and the flexibility of virtue. *Psychological Science*, *18*(8), 689–690.

Van der Lee, R., Ellemers, N., Scheepers, D., & Rutjens, B. T. (2017). In or out? How the perceived morality (vs. competence) of prospective group members affects acceptance and rejection. *European Journal of Social Psychology*, *47*(6), 748–762.

Van Kleef, G. A., Gelfand, M. J., & Jetten, J. (2019). The dynamic nature of social norms: New perspectives on norm development, impact, violation, and enforcement. *Journal of Experimental Social Psychology*, *84*, 103814.

Van Kleef, G. A., Homan, A. C., Finkenauer, C., Gündemir, S., & Stamkou, E. (2011). Breaking the rules to rise to power: How norm violators gain power in the eyes of others. *Social Psychological and Personality Science*, *2*(5), 500–507.

Van Prooijen, A. M., & Ellemers, N. (2015). Does it pay to be moral? How indicators of morality and competence enhance organizational and work team attractiveness. *British Journal of Management*, *26*(2), 225–236.

Van Prooijen, A. M., Ellemers, N., Van der Lee, R., & Scheepers, D. T. (2018). What seems attractive may not always work well: Evaluative and cardiovascular responses to morality and competence levels in decision making teams. *Group Processes & Intergroup Relations*, *21*, 73–87.

Vollhardt, J. R., Cohrs, J. C., Szabó, Z. P., Winiewski, M., Twali, M. S., Hadjiandreou, E., & McNeill, A. (2021). The role of comparative victim beliefs in predicting support for hostile versus prosocial intergroup outcomes. *European Journal of Social Psychology*, *51*(3), 505–524.

Winget, J. R., & Tindale, R. S. (2020). Stereotypic morality: The influence of group membership on moral foundations. *Group Processes & Intergroup Relations*, *23*(5), 710–725.

Yamagishi, T., & Kiyonari, T. (2000). The group as the container of generalized reciprocity. *Social Psychology Quarterly*, *63*(2), 116–132.