

MORAL EMOTIONS ARE RELATIVE

Morality is relative: Anger, disgust and aggression as contingent responses to sibling versus
acquaintance harm

Lukas D. Lopez,* Karie Moorman,* Sara Schneider, Melissa N. Baker, & Colin Holbrook
University of California, Merced

* Authors contributed equally to this manuscript

Author Note

Correspondence concerning this paper should be addressed to Colin Holbrook, Cognitive and
Information Sciences, University of California, Merced, 5200 N. Lake Rd., Merced, CA 95343,
USA. E-mail: cholbrook@ucmerced.edu. All datasets and analysis syntax for these studies have
been archived and may be accessed at: <https://osf.io/8fz6r/>

Abstract

Angry reactions to moral violations should be heightened when wrongs befall oneself in comparison to when wrongs befall acquaintances, as prior research by Molho and colleagues (2017) demonstrates, because aggressive confrontation is inherently risky and therefore only incentivized by natural selection to curtail significant fitness costs. Here, in three pre-registered studies, we extend this sociofunctional perspective to cases of wrongs inflicted on siblings. We observed equivalently heightened anger in response to transgressions against either oneself or one's sibling relative to transgressions against acquaintances across studies, whereas transgressions against acquaintances evoked greater disgust and/or fear (both associated with social avoidance) in two of the three studies. Studies 2 and 3, which incorporated measures of tendencies to confront the transgressor, confirmed that the elevated anger elicited by self or sibling harm partially mediated heightened inclinations toward direct aggression. Finally, in Study 3 we compared tendencies to experience anger and to directly aggress on behalf of siblings and close friends. Despite reporting greater affiliative closeness for friends than for siblings, harm to friends failed to evoke heightened anger relative to acquaintance harm, and participants were inclined to directly aggress against those who had harmed their sibling to a significantly greater extent than when the harm befell their friend. These overall results broadly replicate Molho et al.'s findings and theoretically extend the sociofunctionalist account of moral emotions to kinship.

Keywords: anger; disgust; aggression; siblings; moral cognition; evolutionary psychology

MORAL EMOTIONS ARE RELATIVE

55 Suppose someone tossed a family member's smartphone in a pool for their own
56 amusement. How would you feel? Now imagine that the phone belonged to a stranger. In our
57 lived experience, rage or repugnance often arise unbidden as reflexive responses to immoral acts.
58 From an evolutionary perspective, notwithstanding the idiosyncratic influences of culturally or
59 developmentally contingent factors, our emotional reactions should be calibrated to maximize
60 genetic fitness (Holbrook, forthcoming; Tooby & Cosmides, 2008). Here, we explore the
61 strategic nature of emotional responses within the context of moral transgression.

62 A growing literature highlights anger and disgust as characteristic responses to immoral
63 acts (e.g., Hutcherson & Gross, 2011), although debate continues with regard to whether they
64 truly motivate distinct responses (Nabi, 2002), or why a given individual may experience anger
65 versus disgust in the aftermath of a moral violation (e.g., Russell, Piazza, & Giner-Sorolla,
66 2013). From an adaptationist perspective, the variation in anger versus disgust responses
67 triggered by moral transgressions plausibly derives from the distinct functional outputs of each
68 emotion. Anger motivates tendencies toward directly aggressive confrontation (Carver &
69 Harmon-Jones, 2009), including overt punishment of immoral transgressors (Seip, Van Dijk, &
70 Rotteveel, 2014), to deter future transgressions (Sell, Tooby, & Cosmides, 2009). In contrast to
71 anger, disgust-sensitivity predicts tendencies to avoid the use of violence (Pond et al., 2011).
72 Rather than direct confrontation, disgust motivates "social distancing" from transgressors
73 (Tybur, Lieberman, & Griskevicius, 2009), in a pattern theorized to minimize the likelihood of
74 exposure to such actors and to potentially marshal punishment indirectly via others (Curtis &
75 Biran, 2001).

76 To explore the distinct functional roles postulated for anger versus disgust, Molho and
77 colleagues (2017) experimentally manipulated the fitness costs inflicted by moral transgressors.

MORAL EMOTIONS ARE RELATIVE

They reasoned that, because direct aggression (i.e., verbal or physical confrontation) motivated by anger intrinsically involves risk (i.e., related to physical or social counterattack; Archer & Coyne, 2005), anger responses should be more pronounced when moral violations incur greater fitness costs. Directly confrontational punitive responses are theorized to incentivize transgressors not to repeat their harmful behavior, thereby reducing future fitness costs to a degree sufficient to offset the costs inherent to the risk of confronting the transgressor (Fischer & Roseman, 2007; Sell et al., 2009). Angry responses to relatively uncostly violations would be selected against, as the benefits of reducing minor costs through aggressive deterrence would be outweighed by the substantial possible costs inherent to confrontation. Accordingly, on Molho et al.'s sociofunctionalist account, moral violations that incur fewer costs should tend to arouse disgust, and hence the less immediately effective, yet less costly strategy, of social distancing and indirect retribution (e.g., negative gossip about the transgressor; Archer & Coyne, 2005). To test these predictions, Molho et al. assigned some participants to imagine being personally violated and others to envision the same violations inflicted on an acquaintance (also see Hutcherson & Gross, 2011). As anticipated, they found that immoral harms to the self elicited greater anger, which in turn predicted greater motivation to directly aggress against the transgressor, than did the same harms inflicted on an acquaintance. Conversely, transgressions against acquaintances evoked greater feelings of disgust, and disgust was positively correlated with inclinations to indirectly aggress (for a pre-registered conceptual replication, see Tybur et al., 2019)

Whereas anger appears intrinsically linked with approach-motivation and direct aggression (Harmon-Jones & Allen, 1998), the extent to which disgust is uniquely linked with avoidance-motivation is less obvious, given that fear similarly precipitates withdrawal from

MORAL EMOTIONS ARE RELATIVE

threatening stimuli. Holding aside the evident differences between disgust and fear (discussed below), both emotions are elicited by threat cues and deter risk-taking. Indeed, trait disgust-sensitivity tracks individual differences in risk aversion (Sparks, Fessler, Chan, Ashokkumar, & Holbrook, 2018) and positively correlates with anxiety and tendencies to avoid harm (Olatunji, Armstrong & Elwood, 2017). Accordingly, in the present research on emotional responses to transgressive behavior, it seems plausible that fear would not only significantly correlate with disgust, but also evince a somewhat similar pattern of positive association with indirect, but not direct, aggression toward the transgressor. Like inclinations to feel greater disgust when the costs inflicted are relatively modest, inclinations to feel greater fear would also broadly accord with a sociofunctionalist account to the extent that both disgust and fear deter direct aggression in favor of less risky responses to transgressors.

Despite their coarse thematic similarities with respect to risk-avoidance, disgust and fear also display distinct input-output relationships and thus may not operate equivalently in contexts of moral transgression. For example, experimental exposure to images of pathogen-relevant threats (e.g., feces) trigger not only self-reported disgust but also low-level increases in immune function, whereas exposure to fear-relevant images suggesting violent threat do not (Stevenson et al., 2011). Likewise, disgust-evoking images (e.g., maggots) have been found to sustain attention relative to fear-evoking images (e.g., pointed guns), consistent with the function of disgust in motivating relatively deliberate assessment of contamination risk and appropriate counter-measures, whereas fear of exigent threats tends to motivate rapid responses (Van Hoof, Devue, Vieweg, & Theeues, 2013). Consistent with this interpretation, fear, but not disgust, predicts increases in heart rate (Cisler, Olatunji, & Lohr, 2009). Typical behavioral outcomes of fear, such as freezing, fleeing, or opening the eyes wide to maximize visual awareness, also

MORAL EMOTIONS ARE RELATIVE

qualitatively differ from the behavioral outputs typical of disgust, such as nausea/vomiting, measured withdrawal, and constricting the face to restrict access by pathogens to mouth, nose, and eyes (Tybur et al., 2013). This overall pattern indicates that the threats most evocative of fear (e.g., aggressive conspecifics, precipitous heights) require distinct avoidance tactics relative to the threats most evocative of disgust (Cottrell & Neuberg, 2005; Holbrook, 2016). In light of the functional distinctions between fear and disgust, and the prior literature indicating that feelings of disgust frequently attend cues of moral violations, it may plausibly be the case that, as observed by Molho et al. (2017), disgust evinces a stronger relationship with indirect aggression in response to transgressors than does fear. Given the balance of shared threat-avoidance themes on the one hand, and the individuating input-output logics of the two emotions on the other, the extent to which disgust and fear operate similarly in response to moral transgressions varying in fitness costs is an empirical question difficult to predict on a priori theoretical grounds. We explore this issue in the present research.

The sociofunctional account predicts that moral transgressions which inflict substantial fitness costs should elicit anger. Molho and colleagues manipulated relative cost by comparing reactions to harms to an acquaintance versus the self, but their model makes comparable predictions with regard to harms to other people, provided that said harms entail relatively high fitness costs to the self. *Ceteris paribus*, the heightened costs of harm to kin should also trigger greater anger and direct aggression in contrast to when the same harm befalls acquaintances. Anger and aggression on behalf of another may be conceptualized as a form of aid, and prior research shows that individuals are more willing to incur costs in order to aid close kin than acquaintances, friends, or distant kin (Stewart-Williams, 2007). Beyond genetic fitness benefits, deterring harm to kin via angry intervention also confers benefits insofar as kin typically provide

MORAL EMOTIONS ARE RELATIVE

primary social support in times of need (Burton-Chellew & Dunbar, 2015). Therefore, we tested the sociofunctional model's predictions with regard to immoral harms to siblings as well as to the self.

Hypotheses and Overview of Studies. In three studies, we sought to replicate Molho et al.'s (2017) findings, and to extend the sociofunctional account of anger and disgust to reactions to sibling harm.

We evaluated four interrelated predictions generated by the sociofunctional account of moral emotions:

1) Violations against one's sibling or oneself will evoke greater anger than violations against an acquaintance.

2) Violations against one's sibling or oneself will evoke less disgust than violations against an acquaintance.

3) Violations against one's sibling or oneself will evoke greater inclinations toward direct aggression than violations against an acquaintance.

4) Anger should positively correlate with direct (but not indirect) aggression, whereas disgust should positively correlate with indirect (but not direct) aggression.

Additionally, we exploratorily compared the prevalence and domain-specificity of disgust versus fear reactions in each study.

In Study 1, we examined whether the magnitude of anger and disgust, elicited by imagining five distinct moral violations, is contingent on the identity of the victim: oneself, one's sibling, or an acquaintance. In Study 2, participants again reported their state emotional responses, this time in response to a single moral violation inflicted on the self, a sibling, or an acquaintance, and then rated their inclinations to directly and indirectly aggress against the

MORAL EMOTIONS ARE RELATIVE

transgressor. In Study 3, participants again reported their emotions and aggressive tendencies contingent on the identity of the victim, but with an added target identity condition: a close friend. Study 3 thereby allowed us to test whether a “kinship premium,” hypothesized to adaptively motivate individuals to support family members to a greater extent than is explicable by emotional closeness alone (Curry, Roberts, & Dunbar, 2013), would generalize to responses to moral transgressions inflicted on siblings versus friends, such that participants would experience greater anger, and risk directly aggressing to a greater extent, on behalf of their siblings than on behalf of their friends. In Study 3, we also assessed self-reported emotions using distinct response modes (lexical items and facial arrays), to evaluate the generalizability of the effects of target identity using distinct methods.

The studies in this paper were pre-registered (<https://osf.io/8fz6tr/>), and the full materials, datasets, and analysis syntax are available in the Supplemental Online Material (SOM).¹

Study 1

Methods

Participants. To account for possible overestimations of effect sizes in the original research, and anticipating the need to screen online participants for common issues like incompleteness, we set a large target sample size in Study 1 which, in our three-condition design, works out to approximately 175% per cell of the sample size utilized by Molho and colleagues in each cell of their two-condition version of the same design (Molho et al., 2017, Study 1). We recruited 531 adult participants to complete online surveys using Amazon’s Mechanical Turk platform in exchange for \$0.65 compensation. We screened for incompleteness, failing catch questions, age, and reported sex, yielding a final sample of 465 participants (53.5% male, $M_{age} = 36.79$, $SD = 11.33$).

MORAL EMOTIONS ARE RELATIVE

Procedure. Participants first reported having an adult brother, sister, both, or neither, and were then assigned to one of three conditions (*Self*, $N = 154$; *Sibling*, $N = 145$; *Acquaintance*, $N = 166$). Those with a sibling were randomly assigned to any condition, and others were randomly assigned to either the *Self* or *Acquaintance* conditions. (Follow-up tests confirmed the same overall pattern of results when only including the subset of participants who reported possessing an adult sibling; see SOM.) The target in the *Sibling* condition was described as a sister or brother corresponding with the reported sibling gender, or randomly selected for participants with siblings of both genders. Participants then read five brief scenarios in which the target person is violated (e.g., via theft or deception) following Molho et al. (2017, Study 1).

Participants were next asked to rate the degree to which arrays of faces expressing six emotions (anger, disgust, sadness, surprise, fear, and happiness) corresponded with their own feelings while reading the scenario(s). As in Molho et al.'s design, each array included three distinct male and female faces from the Radboud Faces Database expressing the same emotion (Langner et al., 2010). Facial arrays were employed to circumvent limitations regarding lexical self-report (e.g., linguistic associations between 'anger' and 'disgust') and have been shown to be an effective alternative to lexical items in previous studies of this nature (see Chapman & Anderson, 2013; Gutierrez & Giner-Sorolla, 2007). Participants first selected which one of the six arrays best matched how they felt while reading about the transgression(s), then rated how well each array reflected their feelings according to a 7-point Likert scale (1 = *Strongly disagree*, 7 = *Strongly agree*). Participants were next presented with a forced-choice question probing whether the anger or disgust array best matched their feelings. Finally, participants answered demographic questions before being thanked and debriefed.²

Results

MORAL EMOTIONS ARE RELATIVE

Of the six options, most participants endorsed the anger array (60.4%) or the disgust array (18.9%) as best reflecting their feelings, with relatively low selections of sadness, surprise, fear, or happiness. With regard to participants' mean ratings, anger ($M = 5.70$, $SD = 1.42$) and disgust ($M = 4.46$, $SD = 1.80$) were also most strongly endorsed, with relatively low ratings for sadness, surprise, fear, or happiness (see SOM Table S1). When forced to choose between the anger or disgust arrays, the majority of participants selected anger (78%) over disgust (22%). These patterns notably resemble those reported by Molho et al., 2017, Study 1.

Interaction between target identity and emotion. We tested whether manipulating target identity influenced ratings of anger versus disgust, using a 3 (scenario target; between-subjects) \times 2 (emotion; within-subjects) analysis of variance (ANOVA). Consistent with Prediction 1, the interaction between scenario target and emotion was statistically significant $F(2, 462) = 11.23$, $p < .001$, $\eta_p^2 = .05$ (see Table S3 for descriptives).

Effect of target identity on feelings of anger. A follow-up ANOVA with planned contrasts revealed that, relative to the Acquaintance condition, anger was significantly higher in both the Self condition, $p = .008$, 95% CI = $[-.73, -.11]$, and the Sibling condition, $p = .009$, 95% CI = $[-.74, -.11]$, with no significant differences between the Self and Sibling conditions, $p = .980$. Prediction 1 was therefore supported.

Effect of target identity on feelings of disgust. Consistent with Prediction 2, relative to the Acquaintance condition, ratings of disgust were lower in both the Self condition, $p = .015$, 95% CI = $[-.10, .88]$, and the Sibling condition, $p = .006$, 95% CI = $[-.16, .96]$, with no significant difference between the Self and Sibling, $p = .734$ (see Figure 1). Therefore, Prediction 2 was supported.

MORAL EMOTIONS ARE RELATIVE

Comparisons between disgust and fear. Although few participants selected fear as best reflecting their feelings (2.4%) and the mean levels of fear ($M = 3.31$, $SD = 1.75$) reported in the overall sample were notably lower relative to disgust ($M = 4.46$, $SD = 1.80$), $F(1, 464) = 149.73$, $p < .001$, $\eta_p^2 = .24$, the mean ratings of fear and disgust were positively correlated, $r(464) = .35$, $p < .001$. We tested whether manipulating target identity influenced ratings of anger versus fear, again using a 3 (scenario target; between-subjects) \times 2 (emotion; within-subjects) ANOVA. The interaction between scenario target and emotion was statistically significant, $F(2, 462) = 7.77$, $p < .001$, $\eta_p^2 = .03$, and a follow-up ANOVA with planned contrasts revealed that, relative to the Acquaintance condition, ratings of fear were lower in both the Self condition, $p = .020$, 95% CI = [.07, .84], and the Sibling condition, $p = .041$, 95% CI = [.02, .79], with no significant difference between the Self and Sibling, $p = .804$. Thus, although reported infrequently and at a substantially lower intensity, fear evinced an overall pattern parallel to that of disgust.

Null effects of scenario target on sadness or surprise. Mixed ANOVAs revealed no significant interactions between scenario target and emotion contrasts with anger for mean reported levels of sadness, $p = .871$, or surprise, $p = .052$. However, as there was an apparent marginal interaction with surprise, we conducted a follow-up ANOVA with planned contrasts which confirmed that there were no significant effects of target identity on surprise in the Self condition, $p = .803$, or the Sibling condition, $p = .457$, relative to the Acquaintance condition.

Discussion

In Study 1, we replicated many of the same patterns observed by Molho et al. (2017, Study 1), and extended the sociofunctional account of moral emotions to transgressions against siblings. Hypothetical transgressions against oneself or one's sibling elicited increased anger, and decreased disgust, relative to transgressions against an acquaintance, supporting Predictions

MORAL EMOTIONS ARE RELATIVE

1 and 2. Not only did the sibling target elicit greater anger and reduced disgust relative to the acquaintance, but also closely comparable levels of anger and disgust to when the victim was the self (see Figure 1).

The pattern of relatively increased anger and decreased disgust observed in the self and sibling scenarios is inconsistent with accounts positing that anger and disgust reactions are equivalent. Notably, however, mean levels of self-reported state fear evinced a parallel effect of the target identity manipulation to that of state disgust, consistent with the two emotions' thematic similarity with regard to avoiding threats, and suggesting that in some respects disgust and fear reactions to moral violations may be quite similar. At the same time, disgust was second only to anger as the most frequently selected emotion capturing how participants felt when envisioning the transgressive acts; very few participants selected fear faces as best representing their feelings in regard to transgressions, and the mean intensity of reported fear was markedly lower than the mean intensity of disgust. Thus, similarities in the effects of the scenario manipulation between fear and disgust notwithstanding, disgust appeared to be more relevant to moral transgressions.

Although Study 1 conceptually replicated and extended Molho et al.'s predictions with regard to emotion elicitation, the sociofunctional account focuses on the divergent functions of each emotion in motivating direct versus indirect aggression. Accordingly, Study 2 closely replicated the design of Study 1 (and Molho et al., 2017, Study 4), this time using a transgression scenario featuring a focal adversary regarding whom participants could report their inclinations toward direct and indirect aggression. Study 2 thereby allowed us to test Predictions 3 and 4.

Study 2

Methods

MORAL EMOTIONS ARE RELATIVE

Participants. In Study 2, we lowered our target sample size per cell to approximately 130% of that utilized by Molho et al. (2017, Study 4), as the results of Study 1 indicated that the effect sizes reported by Molho et al. were estimated accurately. We recruited 680 adult participants to complete online surveys on Amazon's Mechanical Turk platform, again in exchange for \$0.65 compensation. We screened according to the same criteria as in Study 1, yielding a final sample of 568 (50.0% male, $M_{age} = 37.73$, $SD = 13.45$).

Procedure. Utilizing the same procedure as in Study 1, participants first reported having an adult brother, sister, both, or neither, and were then assigned to one of the three conditions (*Self*, $N = 188$; *Sibling*, $N = 202$; *Acquaintance*, $N = 178$). (Follow-up tests confirmed the same overall pattern of results when only including the subset of participants who reported possessing an adult sibling; see SOM.) Participants then read one relatively detailed scenario describing a partygoer intentionally ashing his cigarette onto a pile of jackets, ruining a jacket belonging to either the self, a sibling, or an acquaintance (see Molho et al., 2017, Study 4).

Participants were next asked to rate the degree to which arrays of faces expressing six emotions (anger, disgust, sadness, surprise, fear, and happiness) corresponded with their own feelings while reading the scenario(s), as in Study 1. In Study 2, we also assessed inclinations toward direct versus indirect aggression against the transgressor. Following Molho et al., 2017, Study 4, five items measured direct aggression (e.g., "I would hit the person described in the scenario"; $\alpha = .84$) and five items measured indirect aggression (e.g., "I would try to get others to dislike the person described in the scenario"; $\alpha = .88$) according to the same 7-point Likert scale (1 = *Strongly disagree*, 7 = *Strongly agree*). Indirect and direct aggression were significantly correlated, $r(567) = .47$, $p < .001$. The sociofunctional approach does not indicate that direct versus indirect forms of aggression are mutually exclusive, but only that direct [indirect]

MORAL EMOTIONS ARE RELATIVE

strategies should be preferred when the severity of the costs inflicted by a transgressor is relatively high [low]. To elucidate potentially unique relationships between the degree of costs inflicted and inclinations toward direct versus indirect aggression, we planned to control for covariation between the two strategies. Additionally, following Molho et al., we planned to control for participant sex in analyses of aggressive tendencies, given prior research documenting consistent sex differences in direct aggression (Archer, 2004). Finally, participants answered demographic questions before being thanked and debriefed.

Results

Of the six options, most participants endorsed the anger array (53.9%) or the disgust array (22.5%) as best reflecting their feelings, with relatively low selections of sadness, surprise, fear, or happiness. With regard to participants' mean ratings, anger ($M = 5.72$, $SD = 1.42$) and disgust ($M = 4.93$, $SD = 1.78$) were also most strongly endorsed, with relatively low ratings for sadness, surprise, fear, or happiness (see SOM Table S5). When forced to choose between the anger or disgust arrays, the majority of participants selected anger (63.9%) over disgust (36.1%). (These patterns notably resemble those reported in the present Study 1 and by Molho et al.'s closely comparable Study 4.)

Interaction between target identity and emotion. We next tested whether manipulating target identity influenced ratings of anger versus disgust, using a 3 (scenario target; between-subjects) \times 2 (emotion; within-subjects) ANOVA. Consistent with Prediction 1 and the findings of Study 1, the interaction between scenario target and emotion was statistically significant, $F(2, 565) = 7.79$, $p < .001$, $\eta_p^2 = .03$ (see Table S7 for descriptives).

Effect of target identity on feelings of anger. Follow-up ANOVAs with planned contrasts revealed that, relative to the Acquaintance condition, anger was significantly higher in

MORAL EMOTIONS ARE RELATIVE

both the Self condition, $p = .001$, 95% CI = $[-.78, -.20]$, and the Sibling condition, $p = .007$, 95% CI = $[-.67, -.11]$, with no significant differences between the Self and Sibling conditions, $p = .492$ (see Figure 2). Prediction 1 was therefore supported.

Effect of target identity on feelings of disgust. Departing from the results of Study 1, ratings of disgust were not significantly lower in the Self condition, $p = .066$, 95% CI = $[-.02, .71]$ or Sibling condition relative to the Acquaintance condition, $p = .167$, 95% CI = $[-.11, .61]$. Prediction 2 was therefore not supported. As in Study 1, there was no significant difference between the Self and Sibling conditions, $p = .621$ (see Figure 2).

Interaction between target identity and aggression. Effects of target identity on ratings of direct versus indirect aggression were assessed using a 3 (scenario target; between-subjects) \times 2 (aggression; within-subjects) analysis of covariance (ANCOVA; see SOM Table S7 for descriptives). In light of well-documented sex differences in aggression (e.g., Archer, 2004), and following Molho et al. (2017), we controlled for participant sex. Consistent with Prediction 3, the interaction between scenario target and aggression was significant, $F(2, 564) = 7.43$, $p = .001$, $\eta_p^2 = .03$.

Effect of target identity on direct aggression. A follow-up ANCOVA (controlling for sex and covarying preferences for indirect aggression) revealed a main effect of condition on direct aggression, $F(2, 563) = 16.32$, $p < .001$, $\eta_p^2 = .06$. Planned contrasts showed that ratings of direct aggression were significantly lower in the Acquaintance condition relative to both the Self condition, $p < .001$, 95% CI = $[-.96, -.46]$, and Sibling condition, $p < .001$, 95% CI = $[-.75, -.26]$, with no significant difference between the Self and Sibling, $p = .109$ (see Figure 3). Prediction 3 was therefore supported.

MORAL EMOTIONS ARE RELATIVE

Effect of target identity on indirect aggression. An ANCOVA with planned contrasts (controlling for sex and covarying preferences for direct aggression) showed no main effect of condition on ratings of indirect aggression, $p = .135$.

Domain-specific correlations between disgust, anger, and aggression. Consistent with Prediction 4, state anger was correlated with direct aggression, $r(567) = .11$, $p = .007$, but not indirect aggression, $p = .99$, whereas state disgust was correlated with indirect aggression, $r(567) = .09$, $p = .024$, but not direct aggression, $p = .50$.

Anger partially mediates the effect of target identity on direct aggression. Next, we examined whether the significant effects of the Self or Sibling conditions on direct aggression relative to the Acquaintance condition were mediated by anger and/or disgust (entered as simultaneous potential mediators), controlling for participant sex and covarying indirect aggression, using PROCESS for SPSS (Hayes, 2013). Anger was positively related to direct aggression, $b = 0.10$, 95% CI = [0.03, 0.17], $p < .001$, whereas disgust was not, $b = 0.02$, 95% CI = [-0.04, 0.08], $p = .58$. Likewise, the relative direct effects of target condition on direct aggression were significant for both Self, $b = 0.66$, 95% CI = [0.41, 0.92], $p < .001$, and Sibling, $b = 0.47$, 95% CI = [0.23, 0.72], $p < .001$. Lastly, we observed relative indirect effects of target condition on endorsement of direct aggression via anger in the Self condition, $b = 0.05$, and the Sibling condition, $b = 0.04$ (see Figure 4).

Comparisons between disgust and fear. As in Study 1, although few participants selected fear as best reflecting their feelings (6.4%) and the mean levels of fear ($M = 3.89$, $SD = 1.87$) reported in the overall sample were again substantially low relative to disgust ($M = 4.93$, $SD = 1.78$), $F(1, 567) = 128.22$, $p < .001$, $\eta^2 = .18$, mean ratings of fear and disgust were positively correlated, $r(567) = .29$, $p < .001$. We again tested whether manipulating target

MORAL EMOTIONS ARE RELATIVE

identity influenced ratings of anger versus fear, using a 3 (scenario target; between-subjects) \times 2 (emotion; within-subjects) ANOVA. The interaction between scenario target and emotion was statistically significant, $F(2, 565) = 8.66, p < .001, \eta_p^2 = .03$, and follow-up ANOVAs with planned contrasts revealed that, relative to the Acquaintance condition, ratings of fear were lower in both the Self condition, $p = .024$, 95% CI = [.06, .82], and the Sibling condition, $p = .027$, 95% CI = [.05, .80], with no significant difference between the Self and Sibling condition, $p = .93$. Finally, like disgust, mean ratings of fear were positively correlated with indirect aggression, $r(567) = .09, p = .030$, but not direct aggression, $r(567) = -.08, p = .058$. Thus, as in Study 1, although reported infrequently and at a substantially lower intensity, fear again displayed an overall pattern parallel to that of disgust. Notably, fear evinced the effect of the target identity manipulation predicted, but not observed, for disgust.

Null effects of scenario target on sadness and surprise. Departing from expectations and the results of Study 1, a mixed ANOVA revealed a significant interaction between scenario target and emotion contrast with anger for mean reported levels of sadness, $F(2, 565) = 3.97, p = .019, \eta_p^2 = .02$. However, a follow-up ANOVA with planned contrasts confirmed that there were no significant effects of target identity on sadness in the Self condition, $p = .429$, or the Sibling condition, $p = .201$, relative to the Acquaintance condition. A parallel mixed ANOVA revealed a similar interaction for mean reported levels of surprise, $F(2, 565) = 5.62, p = .004, \eta_p^2 = .004$. However, as with sadness, a follow-up ANOVA with planned contrasts revealed no significant effects of target identity on surprise in the Self condition, $p = .089$, or the Sibling condition, $p = .412$, relative to the Acquaintance condition.

Discussion

MORAL EMOTIONS ARE RELATIVE

In Study 2, greater anger was reported when the transgression harmed the self or a sibling relative to an acquaintance, and participants reported greater inclinations to directly aggress against the transgressor when the victim was a sibling or the self, in a pattern that was partially mediated by heightened anger. These findings support Predictions 1, 3, and 4, replicating and extending patterns observed in Study 1 and, with regard to the effects of the Self versus Acquaintance manipulation, the results of Molho et al.'s Study 4 (2017). However, departing from the findings reported by Molho and colleagues, and inconsistent with Prediction 2, we did not find significant effects of the target identity manipulation on either state disgust or indirect aggression.

As in Study 1, state fear displayed a pattern notably parallel to that of state disgust. Both emotions positively correlated with indirect aggression and, strikingly, mean fear was significantly greater when the victim was framed as an acquaintance than when framed as the self or a sibling, in the same pattern that had been predicted to obtain with respect to disgust, but which was not observed in Study 2. As in Study 1, remarkably few participants identified fear as best reflecting their feelings about the transgression relative to disgust, and the mean intensity of reported fear was substantially lower than that of disgust, although the two emotions again appeared to operate comparably in contexts of moral violations. Broadly consistent with a sociofunctional approach, participants reported less fear when fitness costs were high, and greater fear when costs were low and direct confrontation was de-incentivized.

Although our overall pattern of results generally bolsters the sociofunctional account of moral emotions and associated motivations, particularly in regard to anger and direct aggression, both Molho and colleagues (2017, Study 4) and the present Study 2 failed to detect significantly heightened indirect aggression when harm befell acquaintances. Speculatively, the muted effect

MORAL EMOTIONS ARE RELATIVE

of target identity on state disgust observed in Study 2 may owe to the relative mildness of the jacket-ruining transgression scenario in comparison to the violations utilized in Study 1 (e.g., stealing all the money from a bank account). Alternatively, participants may have found it unrealistic to indirectly aggress (e.g., “I would tell a friend an embarrassing secret I’ve heard about this person”) a person at a party, described in the scenario as “a man that you recognize, but whom you’re not friends with,” who they would presumably know little about. Had the scenario been structured in a manner rendering indirect forms of aggression more relevant, such as by framing the transgressor as an acquaintance from within a shared community, depicting the transgression as harming an acquaintance may indeed have significantly heightened reported tendencies toward indirect aggression. These possibilities are explored in Study 3.

Study 3 was also intended to address whether heightened anger and inclinations to directly aggress on behalf of siblings derive from feelings of affiliation which might generalize equally to non-kin, or whether these patterns are token examples of the “kinship premium” hypothesized to adaptively motivate individuals to support family members to a greater extent than is explicable by emotional closeness alone (Curry, Roberts, & Dunbar, 2013). We therefore added a Friend condition, and measures of affiliation applied to either a close adult friend or sibling.

Finally, the preceding two studies and most of Molho et al.’s (2017) research relies on reporting state emotion via facial arrays. Following Molho et al. and other research groups, we have employed facial arrays to address the possibility that lexical self-report measures may not distinguish well between disgust and other negative emotions, particularly anger, due to semantic conflation of the words “anger” and “disgust” among English speakers (Chapman & Anderson, 2013; Nabi, 2002). Self-report using facial arrays provides an alternative to such potential

MORAL EMOTIONS ARE RELATIVE

linguistic confusion, but reliance on any one method raises the possibility that observed effects are bounded by that method. Therefore, Study 3 incorporated lexical measures to assess the generalizability of the results across methods, and to ascertain whether the distinct profiles of anger and disgust become blurred when assessed lexically.

Study 3

Methods

Participants. In Study 3, we recruited 760 adult participants to complete online surveys on Amazon's Mechanical Turk platform in exchange for \$0.65 compensation. The sample size of Study 3 was increased relative to Study 2 due to the addition of a between-subjects Friend condition. We screened according to the same criteria as in Studies 1 and 2, yielding a final sample of 575 (49.4% male, $M_{age} = 38.45$, $SD = 11.73$).

Procedure. Participants first reported having an adult brother, sister, both, or neither and were then assigned to one of four conditions (*Self*, $N = 150$; *Sibling*, $N = 133$; *Friend*, $N = 152$; *Acquaintance*, $N = 140$). Those with a sibling were randomly assigned to any condition, and those without a sibling were randomly assigned to the Friend, Self, or Acquaintance conditions (Follow-up tests confirmed that the same general overall pattern of results when only including the subset of participants who reported possessing an adult sibling; see SOM). Participants then read four brief scenarios in which the target person is violated (e.g., via theft or deception), as in Study 1.³ To portray the transgressor as a member of a shared community with the participant, and thereby render indirect aggression a feasible strategy, the transgressor was described in each scenario as "a guy you know."

Participants were next asked to select which of the four scenarios was most personally upsetting, and then to rate the degree to which the same emotions assessed in the prior two

MORAL EMOTIONS ARE RELATIVE

studies (anger, disgust, sadness, surprise, fear, and happiness) corresponded with their own feelings, while vividly imagining that particular scenario occurring. Emotional responses were rated according to both facial arrays (as in Studies 1 and 2) and lexical terms for all participants (counterbalanced order). The six lexical terms were *angry*, *happy*, *fear*, *grossed out/disgusted*, *surprised*, and *sad*. Participants first selected which one of the six choices (arrays or lexical terms) best matched how they felt while reading about the focal transgression scenario, then were asked to rate how well each array/lexical item reflected their feelings according to a 7-point Likert scale (1 = *Strongly disagree*, 7 = *Strongly agree*). Next, participants were presented with a forced-choice question (answered according to both facial arrays and lexical items) probing whether anger or disgust best matched their feelings. The facial array and lexical results are presented side-by-side to facilitate comparison.

Finally, utilizing the same instruments as in Study 2, we assessed inclinations toward direct aggression ($\alpha = .90$) versus indirect aggression ($\alpha = .86$) against the transgressor. Indirect and direct aggression were again significantly correlated, $r(574) = .51, p < .001$. Accordingly, as in Study 2, we planned to conduct follow-up analyses controlling for covariation between the two strategies in order to reveal potentially unique relationships between experimental condition, emotion, and inclinations toward direct versus indirect aggression.

To explore the role of affiliation, we collected measures of both subjective and objective closeness to the friend or sibling using measures modified from the Adult Sibling Relationship Questionnaire (Lanthier & Stacker, 1992). The overall closeness measure was comprised of two 4-item subscales: subjective closeness (e.g., “How often do you talk to your [sibling/friend] about things that are important to you?”; sibling closeness $\alpha = .92$, friend closeness $\alpha = .94$) and objective closeness (e.g., “How often do you and your [sibling/friend] see each other?”; sibling α

MORAL EMOTIONS ARE RELATIVE

= .96, friend $\alpha = .85$). The two subscales utilized distinct rating scales (objective closeness: 1 = *At least once a week*, 2 = *At least once a month*, 3 = *At least once in 6 months*, 4 = *At least once a year*, 5 = *Less than once a year*; subjective closeness: 1 = *Never*; 2 = *Rarely*; 3 = *Occasionally*; 4 = *Regularly*). In addition, we administered versions of the Relationship Closeness Scale (Dibble, Levine & Park, 2012), which also measures feelings of affiliation, customized to apply to a friend or to a sibling. The scale consisted of ten items (e.g., “When we are apart, I miss my [sibling/friend] a great deal,” “My relationship with my [sibling/friend] is close,” sibling $\alpha = .97$, friend $\alpha = .95$) rated according to a 7-point Likert scale (1 = *Strongly disagree*; 7 = *Strongly agree*). All participants were asked to complete the friend closeness measures. Participants who reported having an adult sibling ($N = 499$) were also asked to complete the sibling closeness measures, yielding a final subsample of 457 participants who completed both sets of measures (see SOM for analyses limited to this subsample). Finally, participants answered demographic questions before being thanked and debriefed.

Results

Of the six facial array options, most participants once again endorsed anger (72.9%) or disgust (13.9%) as best reflecting their feelings, with relatively low selections of sadness, surprise, fear, or happiness. With regard to the six lexical options, most participants also endorsed anger (84.9%), but departing from the pattern observed using facial arrays, more participants selected sadness (6.6%) than disgust (4.7%), potentially because the phrase employed to minimize semantic conflation with ‘anger’ was ‘grossed out / disgusted’. As with the facial arrays, relatively few participants selected lexical surprise or fear as best reflecting their feelings about the transgression.

MORAL EMOTIONS ARE RELATIVE

With regard to participants' mean ratings of each state emotion, anger was most strongly endorsed (*facial*: $M = 6.16$, $SD = 1.06$; *lexical*: $M = 6.46$, $SD = .96$), followed by sadness (*facial*: $M = 4.32$, $SD = 1.77$; *lexical*: $M = 5.01$, $SD = 1.60$) and then disgust (*facial*: $M = 4.30$, $SD = 1.87$; *lexical*: $M = 3.90$, $SD = 1.90$), with relatively low ratings for fear or happiness (see SOM Table S9). Thus, ratings of sadness were unexpectedly evident as an emotional response to the focal moral transgression in Study 3 to a greater extent than in prior studies, in terms of both the facial array and lexical measures, suggesting that when asked to select which scenario was most personally upsetting, participants tended to select a scenario which elicited a relatively high degree of sadness. Also departing from the pattern observed with regard to facial arrays, participants' mean lexical ratings of surprise ($M = 4.40$, $SD = 1.71$) were slightly higher than their lexical ratings of 'grossed-out / disgusted'. When forced to dichotomously choose between anger or disgust, the majority of participants selected anger (*facial*: 84.7%; *lexical*: 95.0%) over disgust (*facial*: 15.3%; *lexical*: 5.0%), as in both prior studies.

Contrasts between Sibling and Friend closeness. We compared feelings of closeness within the subsample of participants who reported having both an adult sibling and a close friend, using a series of within-subjects ANOVAs. With regard to objective closeness, participants reported greater closeness to their friend ($M = 4.00$, $SD = .93$) than to their sibling ($M = 3.20$, $SD = .95$), $F(1, 456) = 142.38$, $p < .001$, $\eta_p^2 = .24$. With regard to subjective closeness, participants also reported greater closeness to their friend ($M = 3.28$, $SD = .73$) than to their sibling ($M = 2.60$, $SD = 1.21$), $F(1, 456) = 101.40$, $p < .001$, $\eta_p^2 = .18$. Finally, participants also reported greater subjective closeness to their friend ($M = 4.98$, $SD = 1.34$) than to their sibling ($M = 4.17$, $SD = 1.74$) when assessed according to the ten-item Relationship Closeness Scale, $F(1, 456) = 85.17$, $p < .001$, $\eta_p^2 = .16$. Follow-up analyses confirmed that this pattern

MORAL EMOTIONS ARE RELATIVE

remains highly significant for all three measures when controlling for target condition, $ps < .001$, within only the subsample of participants assigned to the Sibling condition, $ps < .01$, and within only the subsample of participants assigned to the Friend condition, $ps < .001$. In summary, pooling conditions, participants reported substantially greater objective and subjective closeness to their friends than to their siblings. Follow-up tests also confirmed that including the three paired sibling and friend closeness measures does not alter the overall pattern of results.⁴

Interaction between target identity and emotion. We next tested whether manipulating target identity influenced ratings of anger versus disgust, using a 4 (scenario target; between-subjects) \times 2 (emotion; within-subjects) analysis of variance (ANOVA). Again consistent with Prediction 1 and replicating Studies 1 and 2, the interaction between scenario target and emotion was statistically significant (*facial*: $F[3, 571] = 7.30, p < .001, \eta_p^2 = .04$; *lexical*: $F[3, 571] = 10.41, p < .001, \eta_p^2 = .05$; see Tables S11 and S12 for descriptives).

Effect of target identity on feelings of anger. Follow-up ANOVAs with planned contrasts revealed that, as in Studies 1 and 2, anger was significantly higher in both the Self condition (*facial*: $p = .013, 95\% \text{ CI} = [-.55, -.07]$; *lexical*: $p < .001, 95\% \text{ CI} = [-.67, -.23]$), and the Sibling condition (*facial*: $p = .006, 95\% \text{ CI} = [-.60, -.10]$; *lexical*: $p = .006, 95\% \text{ CI} = [-.54, -.09]$) relative to the Acquaintance condition. Prediction 1 was therefore supported. By contrast, ratings of anger did not significantly differ between the Acquaintance and Friend conditions (*facial*: $p = .254, 95\% \text{ CI} = [-.39, .10]$, *lexical*: $p = .092, 95\% \text{ CI} = [-.03, .40]$) (see Figure 5).

As in Studies 1 and 2, planned contrasts revealed no significant differences between feelings of anger in the Self and Sibling conditions (*facial*: $p = .741$, *lexical*: $p = .215$). However, although there were no significant differences in anger between the Self and the Friend conditions when assessed with facial arrays ($p = .168$), participants did report significantly

MORAL EMOTIONS ARE RELATIVE

higher anger in the Self condition relative to the Friend condition when responding to the lexical item ($p = .015$, 95% CI = [.05, .48]). We detected no significant differences in anger between the Sibling and the Friend conditions (*facial*: $p = .096$, *lexical*: $p = .215$).

Effect of target identity on feelings of disgust. Consistent with Prediction 2 and the results of Study 1, relative to the Acquaintance condition, ratings of disgust were lower in both the Self condition (*facial* [in a marginal effect]: $p = .057$, 95% CI = [-.01, .85]; *lexical*: $p = .003$, 95% CI = [.23, 1.10]), and the Sibling condition (*facial*: $p = .001$, 95% CI = [.32, 1.21]; *lexical*: $p < .001$, 95% CI = [.37, 1.26]). Prediction 2 was therefore supported. Ratings of disgust did not significantly differ between the Acquaintance and the Friend condition when assessed using facial arrays ($p = .338$, 95% CI = [-.22, .64]), although they were significantly lower in the Friend condition when measured with lexical items ($p = .041$, 95% CI = [.02, .88]) (see Figure 5).

Replicating the findings of Studies 1 and 2, planned contrasts revealed no significant differences between feelings of disgust in the Self and Sibling conditions (*facial*: $p = .117$, *lexical*: $p = .505$). Likewise, we observed no significant differences in disgust ratings between the Self and Friend conditions (*facial*: $p = .332$, *lexical*: $p = .323$). However, there were significantly lower levels of disgust reported in the Sibling condition relative to the Friend conditions when assessed with facial arrays ($p = .012$, 95% CI = [-.99, -.12]), but not with lexical items, $p = .104$.

Interaction between target identity and aggression. Effects of target identity on ratings of direct versus indirect aggression were assessed using a 4 (scenario target; between-subjects) \times 2 (aggression; within-subjects) ANCOVA (see SOM Table S11 for descriptives). As in Study 2, we controlled for participant sex. Follow-up tests confirmed that including this covariate did not

MORAL EMOTIONS ARE RELATIVE

alter the pattern of results. Consistent with Prediction 3, and as in Study 2, the interaction between scenario target and aggression was significant, $F(3, 571) = 12.43, p < .001, \eta_p^2 = .06$.

Effect of target identity on direct aggression. A follow-up ANCOVA with planned contrasts (controlling for sex and covarying preferences for indirect aggression) showed that ratings of direct aggression were significantly lower in the Acquaintance condition relative to all three contrast conditions: the Self condition, $p < .001, 95\% \text{ CI} = [-1.44, -.79]$, Sibling condition, $p < .001, 95\% \text{ CI} = [-1.49, -.83]$, and the Friend condition, $p = .004, 95\% \text{ CI} = [-.79, -.16]$. Prediction 3 was therefore supported. Also as observed in Study 2, there was no significant difference in direct aggression ratings between the Self and Sibling conditions, $p = .77$. By contrast, inclinations toward direct aggression were significantly higher in both the Self condition, $p < .001, 95\% \text{ CI} = [-.96, -.33]$, and the Sibling condition, $p = .001, 95\% \text{ CI} = [-1.01, -.37]$, relative to the Friend condition (see Figure 6).

Effect of target identity on indirect aggression. As in Study 2, an ANCOVA with planned contrasts (controlling for sex and covarying preferences for direct aggression) showed no main effect of condition on ratings of indirect aggression, $p = .108$.

Domain-specific correlations between emotion and aggression. As in Study 2, state anger was positively correlated with direct aggression (*facial*: $r[574] = .21, p < .001$; *lexical*: $r[574] = .30, p < .001$). Against expectations, state anger was also positively correlated with indirect aggression (*facial*: $r[574] = .16, p < .001$; *lexical*: $r[574] = .20, p < .001$). We therefore conducted exploratory partial correlations, finding that when controlling for direct aggression, state anger was no longer significantly correlated with indirect aggression (*facial*: $p = .140$; *lexical*: $p = .183$), whereas state anger remained significantly correlated with direct aggression when controlling for indirect aggression (*facial*: $r[572] = .16, p < .001$; *lexical*: $r[572] = .23, p < .001$).

MORAL EMOTIONS ARE RELATIVE

.001). Thus, consistent with Prediction 4, state anger evinced a domain-specific association with direct aggression when covariance with indirect aggression was accounted for. Also consistent with Prediction 4, and as observed in Study 2, state disgust was significantly positively correlated with indirect aggression (*facial*: $r[574] = .12, p = .003$; *lexical*: $r[574] = .09, p = .042$), but not with direct aggression (*facial*: $p = .113$; *lexical*: $p = .173$).

Anger partially mediates the effect of target identity on direct aggression. Next, as in Study 2, we tested whether the significant effects of the Self or Sibling conditions on direct aggression relative to the Acquaintance condition were mediated by anger and/or disgust (entered as simultaneous potential mediators), controlling for participant sex and covarying indirect aggression, using PROCESS for SPSS (Hayes, 2013).

Replicating the results of Study 2, and whether assessed via facial arrays or lexical items, state anger was positively related to direct aggression (*facial*: $b = 0.19, 95\% \text{ CI} = [0.08, 0.29], p < .001$; *lexical*: $b = 0.30, 95\% \text{ CI} = [0.18, 0.42], p < .001$), whereas disgust was not (*facial*: $b = 0.02, 95\% \text{ CI} = [-0.04, 0.08], p = .47$; *lexical*: $b = 0.04, 95\% \text{ CI} = [-0.02, 0.10], p = .21$). The relative direct effects of target condition on direct aggression were significant for both the Self condition (*facial*: $b = 1.03, 95\% \text{ CI} = [0.71, 1.35], p = .005$; *lexical*: $b = 1.11, 95\% \text{ CI} = [0.79, 1.43], p < .001$), Sibling condition (*facial*: $b = 1.11, 95\% \text{ CI} = [0.78, 1.44], p < .001$; *lexical*: $b = 1.16, 95\% \text{ CI} = [0.83, 1.49], p < .001$), and Friend Condition (*facial*: $b = 0.46, 95\% \text{ CI} = [0.14, 0.78], p = .004$; *lexical*: $b = 0.45, 95\% \text{ CI} = [0.13, 0.76], p = .005$). Finally, we observed relative indirect effects of target condition on endorsement of direct aggression via anger in the Self condition (*facial*: $b = 0.05$; *lexical*: $b = .12$), Sibling condition (*facial*: $b = 0.05$; *lexical*: $b = 0.08$), and Friend condition (*facial*: $b = 0.02$; *lexical*: $b = 0.04$) (see Figure 7).

MORAL EMOTIONS ARE RELATIVE

Comparisons between disgust and fear. As in the prior studies, very few participants selected fear as best reflecting their feelings (*facial*: 2.1%; *lexical*: 1.0%) and the mean levels of fear (*facial*: $M = 3.33$, $SD = 1.76$; *lexical*: $M = 2.86$, $SD = 1.74$) were notably lower relative to disgust (*facial*: $M = 4.30$, $SD = 1.87$, $F[1, 574] = 111.51$, $p < .001$, $\eta_p^2 = .16$; *lexical*: $M = 3.90$, $SD = 1.90$, $F[1, 574] = 117.71$, $p < .001$, $\eta_p^2 = .17$). Also as in Studies 1 and 2, mean ratings of fear and disgust were positively correlated (*facial*: $r[574] = .27$, $p < .001$; *lexical*: $r[574] = .22$, $p < .001$). We again tested whether manipulating target identity influenced ratings of anger versus fear, using a 4 (scenario target; between-subjects) \times 2 (emotion; within-subjects) ANOVA. The interaction between scenario target and emotion was not statistically significant when measured using facial arrays, $p = .186$, but it was significant when assessed using lexical items, $F(3, 571) = 5.70$, $p = .001$, $\eta_p^2 = .03$. Follow-up ANOVAs with planned contrasts revealed that, relative to the Acquaintance condition, lexical ratings of fear were higher in both the Self condition, $p < .001$, 95% CI = [-1.61, -.84], and the Sibling condition, $p = .039$, 95% CI = [-.82, -.02], with no such difference in the Friend condition, $p = .227$. This is the reverse of the pattern observed for state disgust. In another departure from the pattern characteristic of disgust ratings, which have not significantly differed between the Self and Sibling conditions in any of the three studies, lexical fear ratings were also significantly higher in the Self condition compared with the Sibling condition, $p < .001$, 95% CI = [-1.20, -.42]. Finally, mean ratings of fear were not significantly positively correlated with indirect aggression when assessed using facial arrays, $p = .195$, but were when assessed with lexical items, $r(574) = .18$, $p < .001$, and this pattern held when controlling for covarying direct aggression. Fear was not correlated with direct aggression when assessed with facial arrays, $p = .761$, but was positively associated with direct aggression when

MORAL EMOTIONS ARE RELATIVE

assessed using lexical items, $r(574) = .13$, $p = .003$, although this relationship did not hold when controlling for covarying tendencies to indirectly aggress, $p = .320$.

Null effects of scenario target on sadness and surprise. Mixed ANOVAs revealed no significant interactions between scenario target and emotion contrasts with anger for mean reported levels of sadness (*facial*: $p = .174$; *lexical*: $p = .682$), or surprise (*facial*: $p = .453$; *lexical*: $p = .218$).

Discussion

Study 3 produced support for all four of the primary predictions, utilizing either facial arrays or lexical items to measure state emotion. Violations framed as harming either a sibling or the self elicited greater anger ratings (Prediction 1), lower disgust ratings (Prediction 2), and stronger inclinations toward direct aggression (Prediction 3) than when framed as harming an acquaintance. Consistent with Prediction 4, anger was positively associated with direct (but not indirect) aggression, with the reverse pattern observed for disgust.

The primary objective of Study 3 was to assess whether immoral acts of harm to friends versus siblings would comparably arouse anger, and related tendencies toward direct aggression, to a greater extent than when harm befalls acquaintances, or whether relatedness would advantage such reactions on behalf of siblings in line with the kinship premium hypothesis (Curry, Roberts, & Dunbar, 2013). We found that participants reported feeling dramatically closer to their friends than to their siblings when assessed via three distinct measures. Nevertheless—and somewhat surprisingly—harm to friends aroused comparable anger to harm inflicted on a mild acquaintance, and harm to siblings aroused significantly more direct aggression than did harm to friends. This overall pattern suggests that a kinship premium may indeed potentiate greater anger and direct aggression when siblings are transgressed against,

MORAL EMOTIONS ARE RELATIVE

plausibly reflecting an ultimate incentive to deter harm to kin that functions somewhat orthogonally to proximate mechanisms of affiliation. However, the magnitude of the differences in reactions to transgressions against siblings versus friends should not be exaggerated. Although harm to a sibling, but not a friend, evoked significantly greater anger than harm to an acquaintance, the mean anger ratings between the Sibling and Friend conditions were not significantly different. Likewise, although the sibling manipulation caused significantly greater direct aggression ratings than the friend manipulation, it should be noted that, in line with recent research on third party punishment (Pedersen, McAuliffe, & McCullough, 2018), participants were more willing to directly aggress on behalf of their friends than on behalf of a mere acquaintance (see Figure 6).

The parallels between disgust and fear observed in our previous studies were not evident in Study 3. Diverging from the findings of Studies 1 and 2, there were no effects of the scenario target manipulation on fear comparable to the effects observed for disgust when measured with facial arrays. Moreover, the lexical measure actually indicated increased fear in the Self and Sibling conditions, whereas disgust significantly decreased in those conditions whether measured lexically or with facial arrays. As the comparisons between disgust and fear have been exploratory in nature, and as the discrepant results of Study 3 were unexpected and evident only in the lexical measure, we will not engage in further conjecture on the matter. At present, however, the findings of Study 3 can be taken as evidence that the effects of victim identity on fear and disgust appear to diverge in some contexts.

Finally, the state affect findings obtained using facial array ratings in the previous studies appeared mostly to generalize when using lexical measures. We observed comparable results of the scenario target identity manipulation ratings of both anger and disgust, as well as comparable

MORAL EMOTIONS ARE RELATIVE

patterns of significant correlations between anger, disgust, and direct or indirect aggression. Speculatively, our choice of the phrase “grossed out / disgusted” rather than “disgusted” may have helped to avert semantic conflation of ‘disgust’ and ‘anger’ by highlighting the nausea component specific to disgust.⁵ Alongside the broadly equivalent results obtained with the facial and lexical arrays, there were notable differences in the results obtained with the two methodological modalities in regard to the percentages of participants selecting which emotion best reflected their feelings. Thus, while the overall generalizability of the effects provide reassurance that the present findings and those of Molho and colleagues (2017) are not mere methodological artifacts of the use of facial arrays to rate state emotion, the differences we did observe motivate some caution, and continued use of convergent methods in future studies.

General Discussion

In three studies, we sought to replicate and extend Molho et al.’s (2017) account of moral emotions to transgressions against siblings. Consistent with the sociofunctional approach, hypothetical transgressions against oneself or one’s sibling reliably elicited heightened anger and inclinations toward direct aggression relative to transgressions against an acquaintance (supporting Predictions 1 and 3). Also echoing Molho et al.’s findings, transgressions against an acquaintance elicited greater disgust than transgressions against oneself or one’s sibling in Study 1 and Study 3 (supporting Prediction 2), with similar trends observed in Study 2. In further support of the sociofunctional model, anger consistently predicted direct (but not indirect) aggression, while disgust evinced the reverse pattern (supporting Prediction 4). Further bolstering the functional specificity of disgust (and to some extent fear) relative to anger, we observed no comparable effects of condition on sadness or surprise.

MORAL EMOTIONS ARE RELATIVE

Although in broad outline the findings of the present studies accord with the sociofunctional hypothesis, there are empirical inconsistencies worth noting. In Studies 2 and 3 (as well as in Molho and colleagues' [2017] Study 4), manipulating target identity did not significantly influence tendencies toward indirect aggression, despite the effort we made in modifying the design of Study 3 to render indirect aggression a seemingly viable option. At the level of method, inclinations toward direct aggression may be easier to manipulate because directly aggressive strategies readily apply to transgressing strangers, whereas counterfactually imagining acquaintanceship with a fictional transgressor embedded in a quasi-fictional shared community may place problematic representational demands on participants that render the prospect of indirect aggression less salient. At the level of theory, a simpler interpretation may be that the evolved psychology is more attuned to factors that incentivize versus de-incentivize direct aggression than indirect aggression, as direct aggression carries greater potential fitness costs (i.e., physical or reputational harm) and payoffs (i.e., deterrence of future transgression).

In another set of findings somewhat at odds with Molho et al.'s portrayal of disgust as deterring direct aggression in a domain-specific manner, mean levels of self-reported state fear evinced a pattern parallel to that of state disgust in Studies 1 and 2 (but not Study 3). In both of these studies, state fear was significantly higher in the Acquaintance condition than in either the Self or Sibling conditions, and fear was also significantly associated with indirect aggression, but not direct aggression. On the one hand, these parallel effects of fear may be taken as evidence against the specialization of disgust responses to moral transgressions. On the other hand, a relatively minute proportion of participants in any of the three studies selected fear faces as best matching their feelings in response to the transgression scenarios (2.1% - 6.4% across studies) in comparison with the number of participants who selected disgust faces (13.9% - 22.5% across

MORAL EMOTIONS ARE RELATIVE

studies), and participants consistently reported a markedly greater mean intensity of feelings of disgust relative to fear.⁶ Thus, while it is sensible that individuals would experience fear in response to transgressive acts of harm, the present data agree with the sizable prior literature showing that anger and disgust are the predominant emotional responses to moral transgressions.

The observed link between fear and indirect—but not direct—aggression is also compatible with a sociofunctional approach to aggressive responses to moral transgressions for two reasons. First, to the extent that fear of the transgressive act indexes the perceived risk posed by the transgressor, directly aggressive behavioral responses should be discouraged. Second, participants reported greater fear in response to transgressions against an acquaintance than against the self or a sibling, whereas an account of emotional responses to moral transgressions which equates emotions of negative valence would predict the reverse pattern, such that harms to the self or kin would elicit greater fear (as well as greater disgust). By contrast, the sociofunctional account linking particular emotions to particular motivations can make sense of why participants would feel less fear (yet more anger and direct aggression) when the self or a sibling are harmed than when a near-stranger is harmed.

Study 3 went beyond the aim of replicating and extending Molho and colleagues' work to explore the potential proximate mechanisms through which harm to siblings evokes greater anger, greater direct aggression, and less disgust than harm to acquaintances. Could emotional affiliation mediate these differences? The kinship premium hypothesis contends that close kin will be helped to a greater extent than is explained by the tendency to support others due to emotional closeness (Curry, Roberts, & Dunbar, 2013). Importantly, the sociofunctional perspective does not entail that such a kinship premium exist, but rather that the risks inherent to anger and direct aggression should be taken in response to relatively costly transgressions of any

MORAL EMOTIONS ARE RELATIVE

kind. Transgressions against friends can potentially inflict fitness costs to the self by limiting the friend's future capacity to aid the self, in addition to potential costs related to damage to one's reputation and/or the friendship in the event that one does not aid one's friend in confronting their transgressor. Thus, the sociofunctional account does not suggest that individuals would display muted patterns of emotional or aggressive responses to friend harm in comparison to sibling harm, nor does the sociofunctional account conflict with this possibility, largely because no role for proximate feelings of affiliation has been theoretically specified. The potential existence of a kinship premium in the context of responses to moral transgressions was therefore an open empirical question. In the event, the overall findings of Study 3 unambiguously favored the psychological reality of a kinship premium in responses to immoral harm. These results were exploratory in nature and invite replication, but can be advanced for now as preliminary evidence that the parallel patterns observed with regard to responses to harm to self and sibling are not driven by feelings of kin-affiliation.

Conclusion

In a series of pre-registered studies, we repeatedly replicated and extended prior work rooted in the premise that natural selection has shaped emotional reactions to moral transgressions to contingently track the fitness incentives of direct aggression. It should come as little surprise that Molho et al.'s (2017) findings proved generally robust, given that their original studies were also publicly archived, and have been largely conceptually replicated in a recent pre-registered study (Tybur et al., 2019). Replications which incorporate straightforward theoretical extensions, as we have pursued here, hold particular promise in building a cumulative science. It is our hope that other emotion researchers will similarly pursue replication-and-extension projects, not only of studies which are dubious by dint of lacking pre-registration, but

MORAL EMOTIONS ARE RELATIVE

784 also of theoretically cogent, methodologically transparent work conducted in the spirit of open
785 science. Replicable results are primarily valuable to the extent that they capture phenomena
786 which translate to the real world. Therefore, beyond establishing replicability, researchers should
787 also take steps to establish validity outside of laboratory or online environments. In the case of
788 the present results, for example, objective behavioral measures of aggression should be
789 employed in place of hypothetical questions to reveal whether kinship truly determines the extent
790 to which individuals directly confront transgressors. Our successful replication and extension of
791 Molho et al.'s sociofunctional account justifies investment in such behavioral research efforts,
792 and provides a novel lens—kinship—through which to observe the strategically contingent
793 nature of moral emotions.

References

- Archer, J. (2004). Sex differences in aggression in real-world settings: A meta-analytic review. *Review of General Psychology*, 8, 291–322. doi:10.1037/1089-2680.8.4.291
- Archer, J., & Coyne, S. M. (2005). An integrated review of indirect, relational, and social aggression. *Personality and Social Psychology Review*, 9, 212–230. doi:10.1207/s15327957pspr0903_2
- Burton-Chellew, M. N., & Dunbar, R. I. M. (2015). Hamilton's rule predicts anticipated social support in humans. *Behavioral Ecology*, 26, 130-137. doi:10.1093/beheco/aru165
- Carver, C. S., & Harmon-Jones, E. (2009). Anger is an approach related affect: Evidence and implications. *Psychological Bulletin*, 135, 183–204. doi:10.1037/a0013965
- Chapman, H.A. & Anderson, A.K. (2013). Things rank and gross in nature: A review and synthesis of moral disgust. *Psychological Bulletin*, 139, 300-327.
- Cisler, J. M., Olatunji, B. O., & Lohr, J. M. (2009). Disgust, fear, and the anxiety disorders: A critical review. *Clinical Psychology Review*, 29, 34-46.
- Cottrell, C. A., & Neuberg, S. L. (2005). Different emotional reactions to different groups: a sociofunctional threat-based approach to "prejudice". *Journal of Personality and Social Psychology*, 88, 770.
- Curry, O., Roberts, S. G. B., & Dunbar, R. I. M. (2013). Altruism in social networks: Evidence for a 'kinship premium'. *British Journal of Psychology*, 104, 2. doi:10.1111/j.2044-8295.2012.02119.x
- Curtis, V., & Biran, A. (2001). Dirt, disgust, and disease: Is hygiene in our genes? *Perspectives in Biology and Medicine*, 44, 17–31. doi:10.1353/pbm.2001.0001
- Dibble, J. L., Levine, T. R., & Park, H. (2012). The Unidimensional Relationship Closeness

MORAL EMOTIONS ARE RELATIVE

- 817 Scale (URCS): Reliability and validity evidence for a new measure of relationship
 818 closeness. *Psychological Assessment*, 24, 565-572.
- 819 Fischer, A. H., & Roseman, I. J. (2007). Beat them or ban them: The characteristics and social
 820 functions of anger and contempt. *Journal of Personality and Social Psychology*, 93,
 821 103–115. doi:10.1037/0022-3514.93.1.103
- 822 Gutierrez, R., & Giner-Sorolla, R. (2007). Anger, disgust, and presumption of harm as reactions
 823 to taboo-breaking behaviors. *Emotion*, 7, 853-868.
- 824 Hayes, A. F. (2013). Model templates for PROCESS for SPSS and SAS.
- 825 Harmon-Jones, E., & Allen, J. J. B. (1998). Anger and frontal brain activity: EEG asymmetry
 826 consistent with approach motivation despite negative affective valence. *Journal of*
 827 *Personality and Social Psychology*, 74, 1310–1316. doi:10.1037/0022-3514.74.5.1310
- 828 Hutcherson, C. A., & Gross, J. J. (2011). The moral emotions: A social-functionalist account of
 829 anger, disgust, and contempt. *Journal of Personality and Social Psychology*, 100, 719–
 830 737. doi:10.1037/a0022408
- 831 Holbrook, C. (forthcoming). Adaptation, co-optation, variation: Emotions in evolutionary
 832 context. In A. Scarantino, (Ed.), *Routledge Handbook of Emotion Theory*. New York:
 833 Taylor & Francis.
- 834 Holbrook, C. (2016). Branches of a twisting tree: Domain-specific threat psychologies derive
 835 from shared mechanisms. *Current Opinion in Psychology*, 7, 81-86.
 836 doi.org/10.1016/j.copsyc.2015.08.006
- 837 Langner, O., Dotsch, R., Bijlstra, G., Wigboldus, D.H.J., Hawk, S.T., & van Knippenberg, A.
 838 (2010). Presentation and validation of the Radboud Faces Database. *Cognition &*
 839 *Emotion*, 24, 1377-1388.

MORAL EMOTIONS ARE RELATIVE

- 840 Lanthier, R.P., & Stacker, C. (1992). *The Adult Sibling Relationship Questionnaire*. Denver,
841 CO: University of Denver.
- 842 Molho, C., Tybur, J. M., Güler, E., Balliet, D., & Hofmann, W. (2017). Disgust and anger relate
843 to different aggressive responses to moral violations. *Psychological Science*, 28, 609-619.
- 844 Nabi, R. L. (2002). The theoretical versus the lay meaning of disgust: Implications for emotion
845 research. *Cognition & Emotion*, 16, 695–703. doi:10.1080/02699930143000437
- 846 Olatunji, B. O., Armstrong, T., & Elwood, L. (2017). Is disgust proneness associated with
847 anxiety and related disorders? A qualitative review and meta-analysis of group
848 comparison and correlational studies. *Perspectives on Psychological Science*, 12, 613-
849 648.
- 850 Pedersen, E. J., McAuliffe, W. H., & McCullough, M. E. (2018). The unresponsive avenger:
851 More evidence that disinterested third parties do not punish altruistically. *Journal of*
852 *Experimental Psychology: General*, 147, 514.
- 853 Pond Jr, R. S., DeWall, C. N., Lambert, N. M., Deckman, T., Bonser, I. M., & Fincham, F. D.
854 (2012). Repulsed by violence: disgust sensitivity buffers trait, behavioral, and daily
855 aggression. *Journal of Personality and Social Psychology*, 102, 175.
- 856 Russell, P. S., Piazza, J., & Giner-Sorolla, R. (2013). CAD revisited: Effects of the word moral
857 on the moral relevance of disgust (and other emotions). *Social Psychological &*
858 *Personality Science*, 4, 62–68. doi:10.1177/1948550612442913
- 859 Seip, E. C., Van Dijk, W. W., & Rotteveel, M. (2014). Anger motivates costly punishment of
860 unfair behavior. *Motivation and Emotion*, 38, 578–588. doi:10.1007/s11031-014-9395-4
- 861 Sell, A., Tooby, J., & Cosmides, L. (2009). Formidability and the logic of human anger.
862 *Proceedings of the National Academy of Sciences, USA*, 106, 15073–15078.

MORAL EMOTIONS ARE RELATIVE

- 863 doi:10.1073/pnas.0904312106
- 864 Sparks, A., Fessler, D.M.T., Chan, K.Q., Ashokkumar, A., & Holbrook, C. (2018). Disgust as a
 865 mechanism for decision making under risk: Illuminating sex differences and individual
 866 risk-taking correlates of disgust propensity. *Emotion, 18*, 942-958.
- 867 Stewart-Williams, S. (2007). Altruism among kin vs. nonkin: effects of cost of help and
 868 reciprocal exchange. *Evolution and Human Behavior, 28*, 193-198. doi:10.1016
- 869 Tooby, J., & Cosmides, L. (2008). The evolutionary psychology of the emotions and their
 870 relationship to internal regulatory variables. Lewis, M., Haviland-Jones, J. M., & Barrett,
 871 L. F. (2010). *Handbook of emotions*. New York: Guilford Press.
- 872 Tybur, J. M., Lieberman, D., & Griskevicius, V. (2009). Microbes, mating, and morality:
 873 Individual differences in three functional domains of disgust. *Journal of Personality and*
 874 *Social Psychology, 97*, 103–122. doi:10.1037/a0015474
- 875 Tybur, J. M., Lieberman, D., Kurzban, R., & DeScioli, P. (2013). Disgust: Evolved function and
 876 structure. *Psychological Review, 120*, 65.
- 877 Tybur, J. M., Molho, C., Dores Cruz, T. D., Cakmak, B., Singh, G. D., & Zwicker, M. (2019,
 878 January 29). Disgust, Anger, and Aggression: Further Tests of the Equivalence of Moral
 879 Emotions. <https://doi.org/10.31234/osf.io/wuspt>
- 880 Van Hooff, J. C., Devue, C., Vieweg, P. E., & Theeuwes, J. (2013). Disgust-and not fear-
 881 evoking images hold our attention. *Acta Psychologica, 143*, 1-6.
- 882

MORAL EMOTIONS ARE RELATIVE

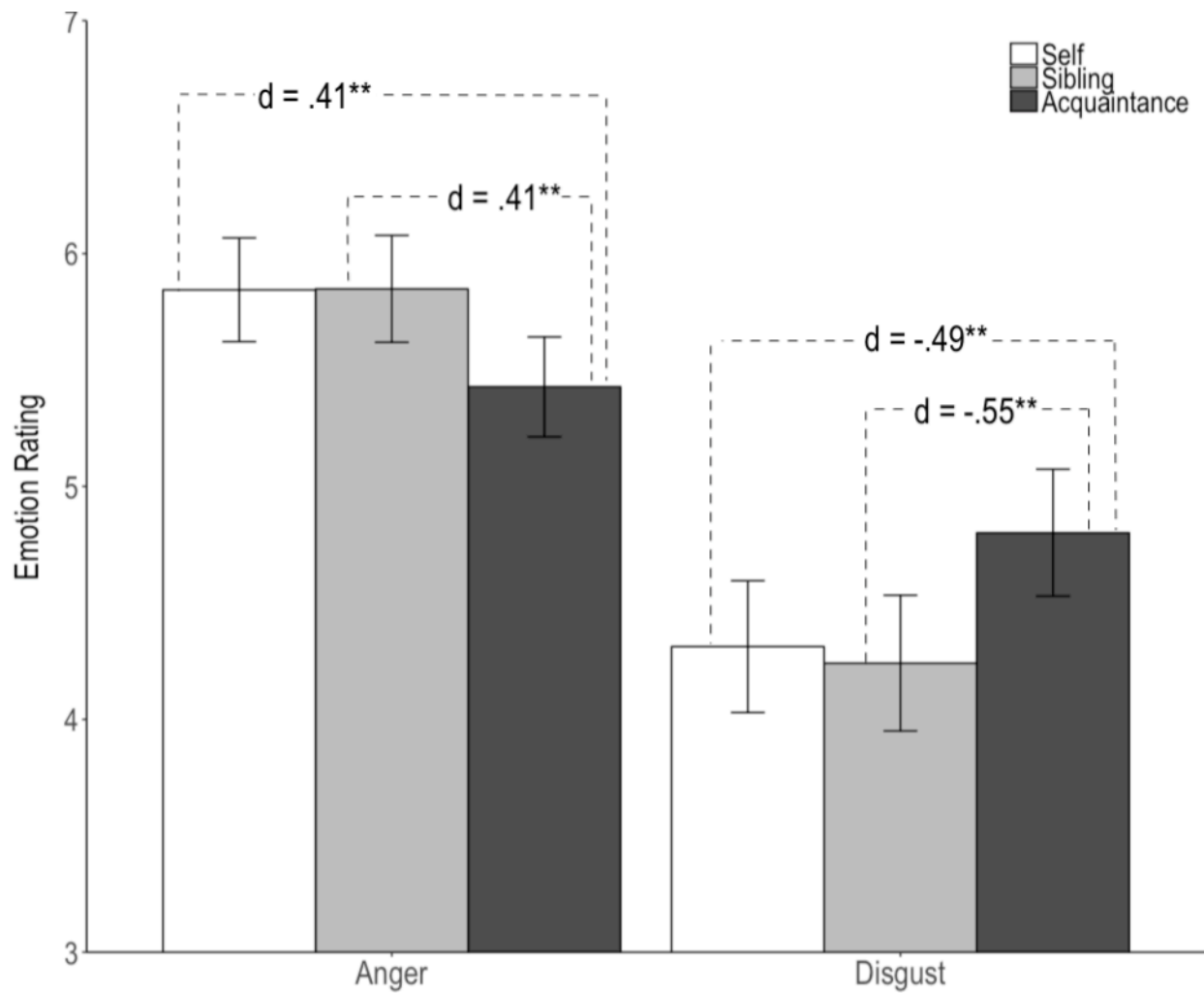


Figure 1. Mean ratings of anger and disgust by target condition in Study. Error bars indicate 95% confidence intervals. The dagger and asterisks indicate the significance of the differences between conditions ($\dagger p < .10$, $*p < .05$, $**p < .01$, $***p < .001$).

MORAL EMOTIONS ARE RELATIVE

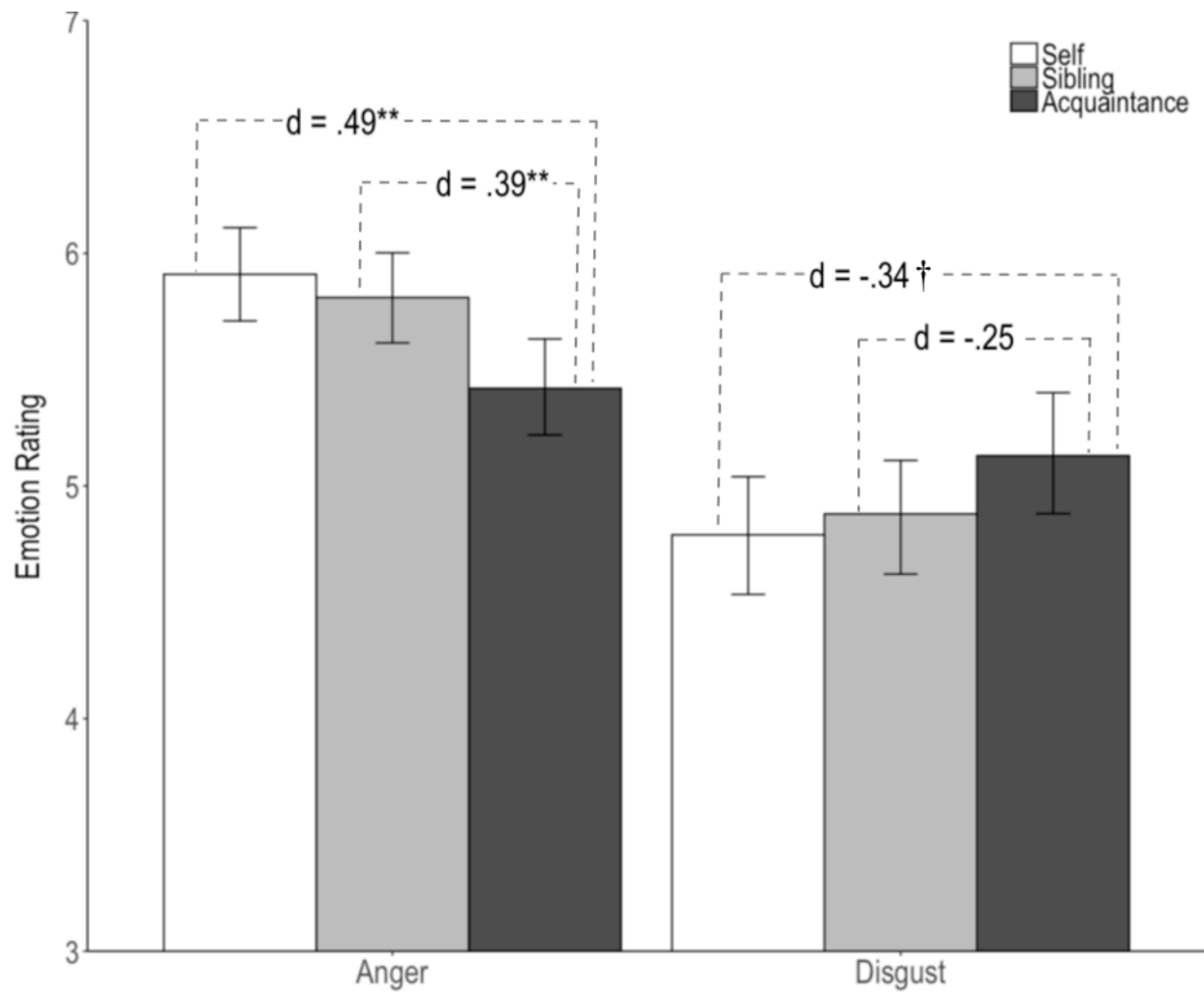


Figure 2. Mean ratings of anger and disgust by target condition in Study 2. Error bars indicate 95% confidence intervals. The dagger and asterisks indicate the significance of the differences between conditions ($\dagger p < .10$, $*p < .05$, $**p < .01$, $***p < .001$).

MORAL EMOTIONS ARE RELATIVE

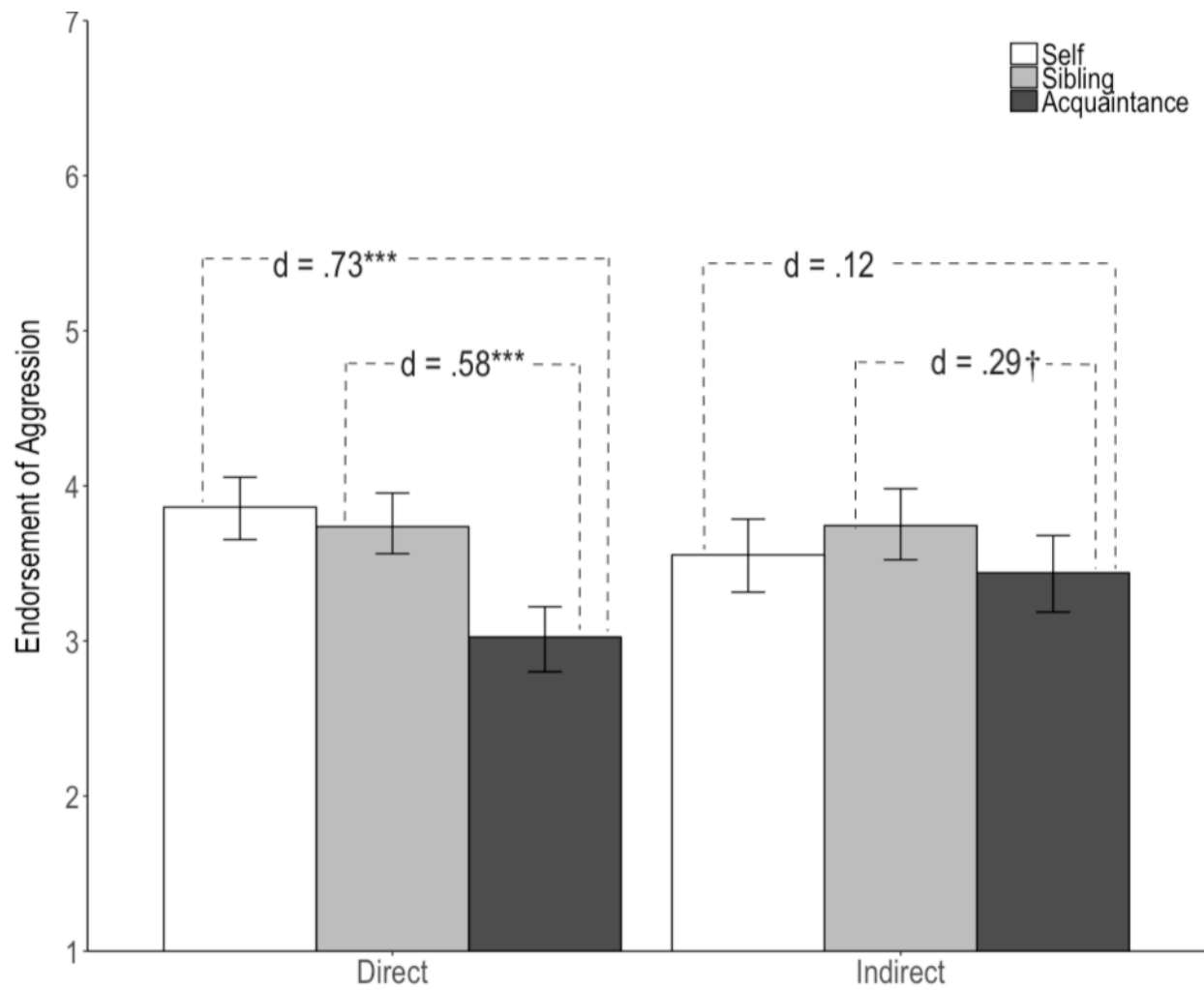


Figure 3. Mean ratings of direct and indirect aggression by target condition in Study 2. Error bars indicate 95% confidence intervals. The dagger and asterisks indicate the significance of the differences between conditions ($^{\dagger}p < .10$, $*p < .05$, $**p < .01$, $***p < .001$).

MORAL EMOTIONS ARE RELATIVE

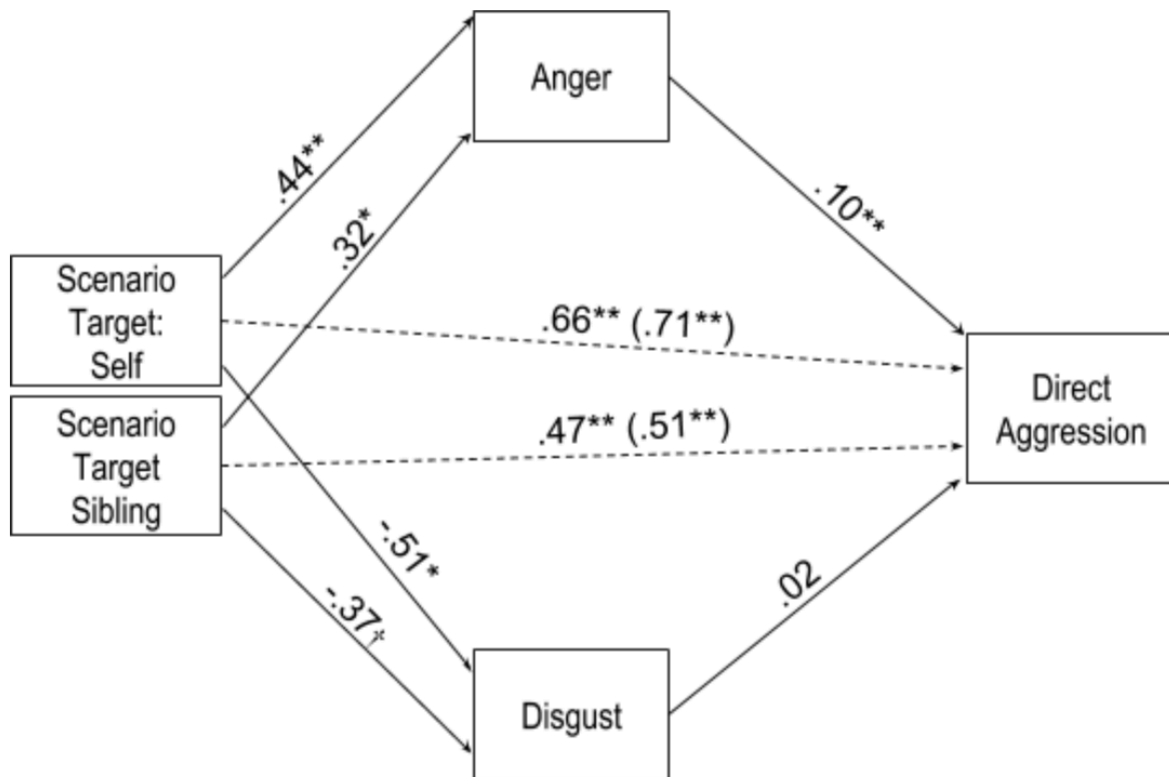


Figure 4. Unstandardized regression coefficients for the effects of scenario targets on endorsement of direct aggression, as mediated by ratings of anger and disgust (Study 2). The models controlled for participant sex, covariances between anger and disgust ratings, and endorsement of indirect aggression. Values above the dashed arrows refer to residual and total (in parentheses) direct effects of target identity on endorsements of aggression. Values above the solid arrows refer to the effects of scenario targets on the potential mediators (anger and disgust) and their effects on endorsements of direct aggression. The dagger and asterisks indicate marginally significant and significant paths ($†p < .10$, $*p < .05$, $**p < .01$).

MORAL EMOTIONS ARE RELATIVE

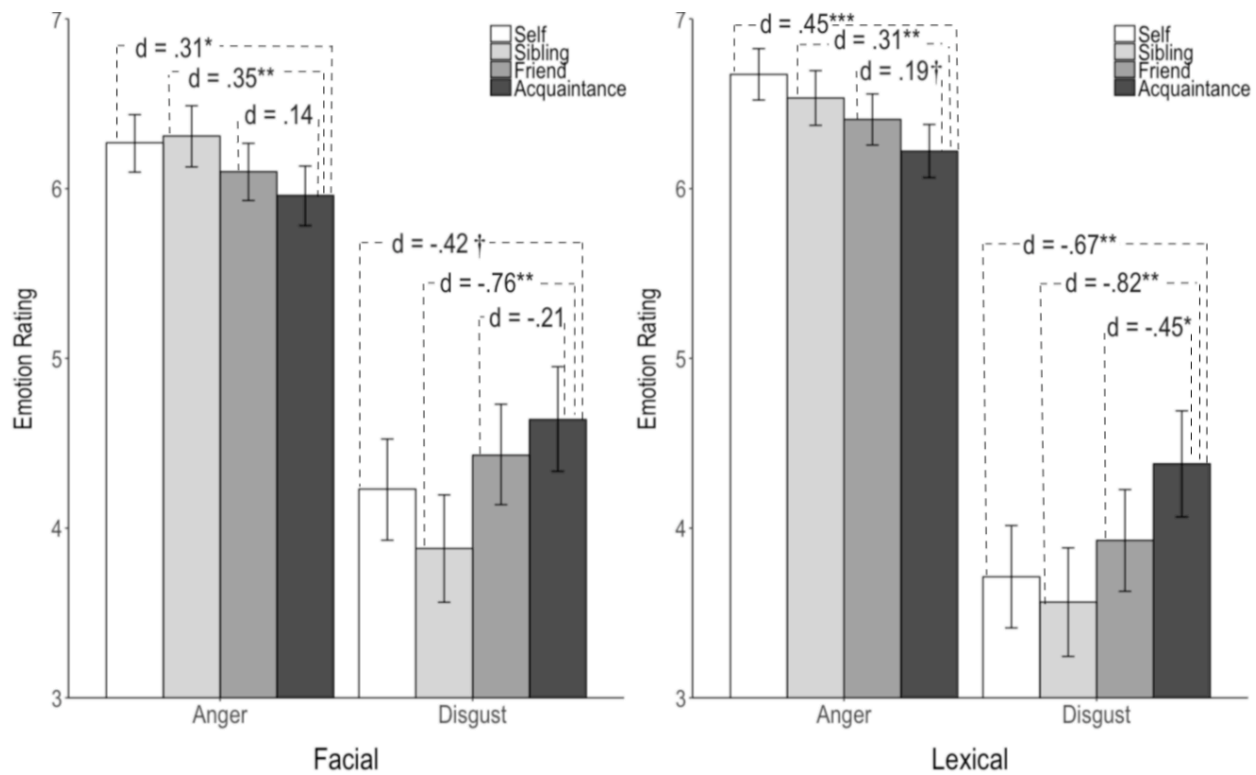


Figure 5. Mean ratings of anger and disgust by target condition in Study 3 (A) and Study 2 (B). Error bars indicate 95% confidence intervals. The dagger and asterisks indicate the significance of the differences between conditions ($^\dagger p < .10$, $*p < .05$, $**p < .01$, $***p < .001$).

MORAL EMOTIONS ARE RELATIVE

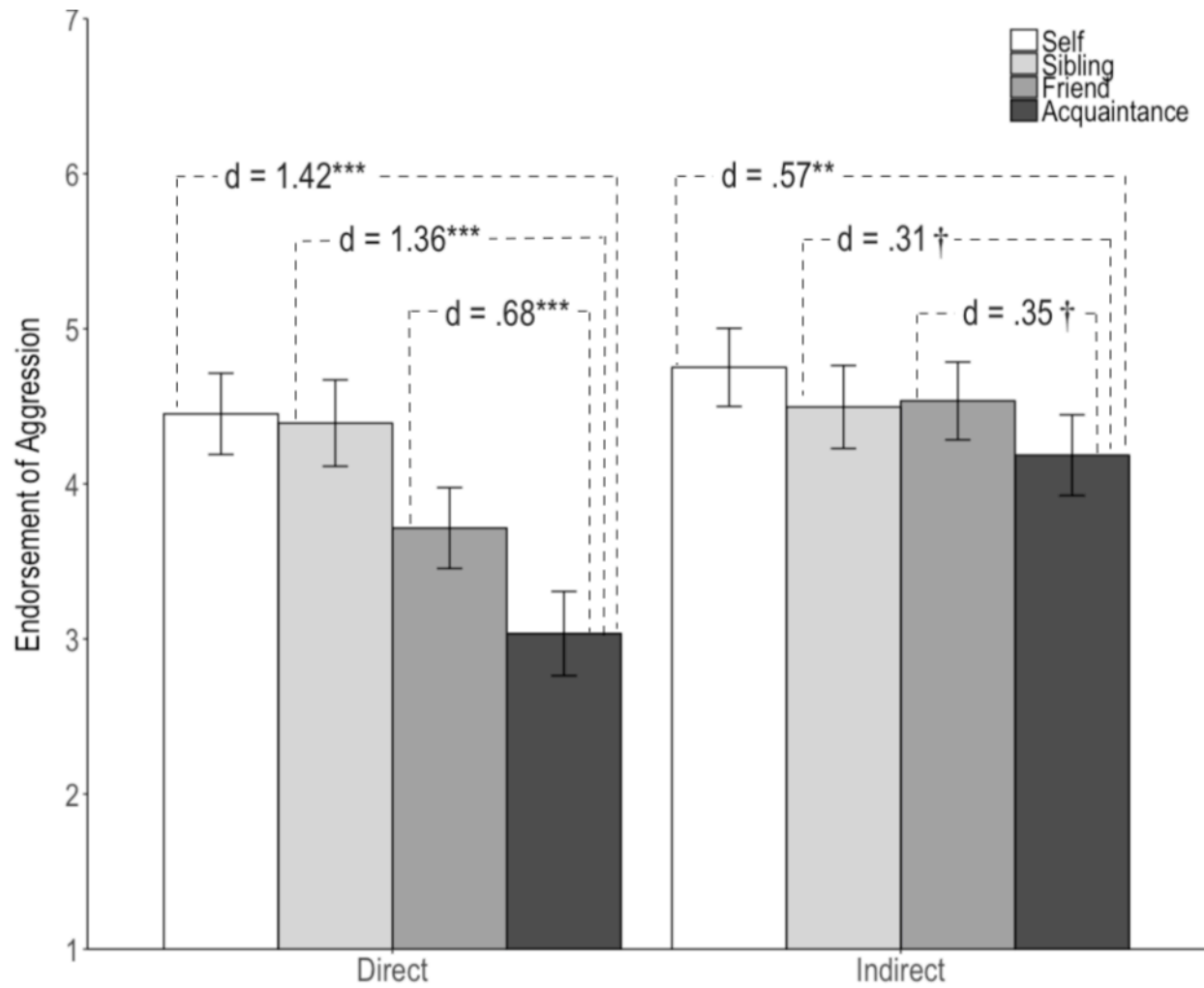
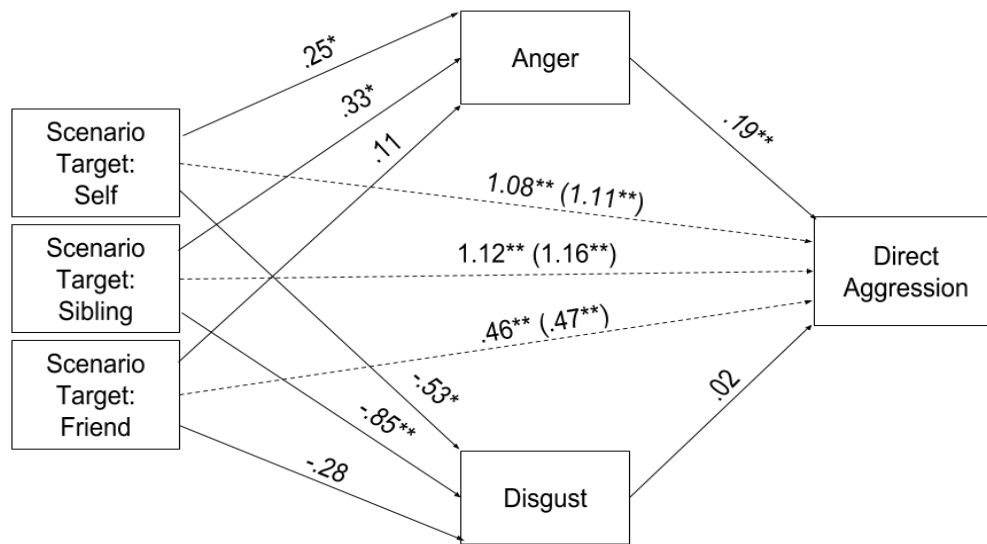
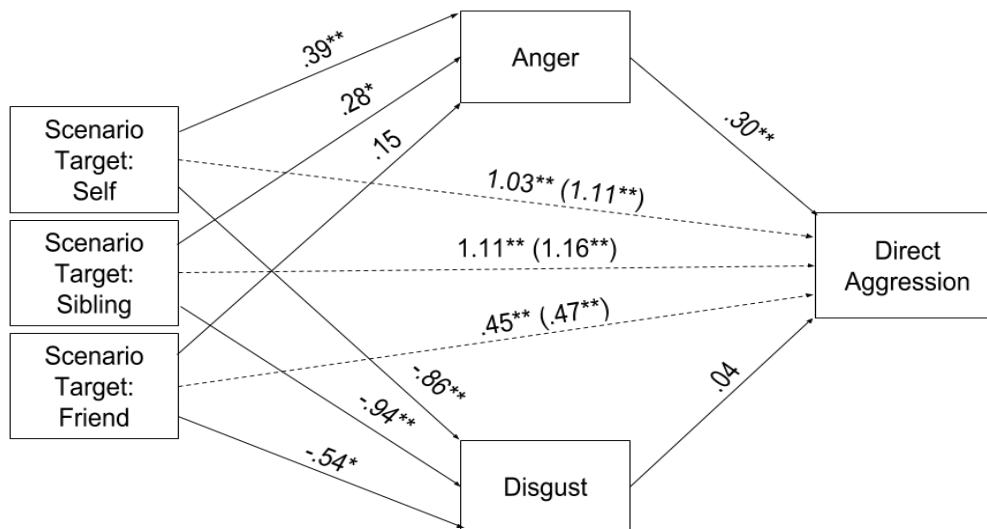


Figure 6. Mean ratings of direct and indirect aggression by target condition in Study 3. Error bars indicate 95% confidence intervals. The dagger and asterisks indicate the significance of the differences between conditions ($^{\dagger}p < .10$, $*p < .05$, $**p < .01$, $***p < .001$).

MORAL EMOTIONS ARE RELATIVE



Facial



Lexical

Figure 7. Unstandardized regression coefficients for the effects of scenario targets on endorsement of direct aggression, as mediated by ratings of anger and disgust (Study 3). The models controlled for participant sex, covariances between anger and disgust ratings, and endorsement of indirect aggression. Values above the dashed arrows refer to residual and total

MORAL EMOTIONS ARE RELATIVE

(in parentheses) direct effects of target identity on endorsements of aggression. Values above the solid arrows refer to the effects of scenario targets on the potential mediators (anger and disgust) and their effects on endorsements of direct aggression. The dagger and asterisks indicate marginally significant and significant paths ($\dagger p < .10$, $*p < .05$, $**p < .01$).

MORAL EMOTIONS ARE RELATIVE

Footnotes

¹ In the course of research on potential links between moral emotions and political orientation, we conducted a similar experiment to Study 1, but in which MTurk participants were recruited on the basis of a party affiliation criterion, yielding a similar but muted effect of target identity. Although the findings related to political orientation are withheld pending separate publication, the results of the measures that were identical to Study 1 are provided in the SOM.

² In all three studies, we also measured trait differences in aggression, disgust and sibling closeness to investigate potential interactions with target identity and state emotion in exploratory analyses (pre-registered as such). Controlling for these variables does not alter the pattern of results (see SOM).

³ One of the five scenarios used in Study 1 and by Molho and colleagues (2017, Study 1) involved someone sleeping with a romantic partner. This scenario would lead female participants, in many cases, to imagine the transgressor as female. Accordingly, we did not include this scenario in Study 3, in order to keep the imagined sex of the transgressor male, as intended, to parallel Study 2.

⁴ Although the overall pattern is not altered when controlling for closeness, there is one outcome which does change. When controlling for objective closeness, subjective closeness, or subjective closeness measured with the Relationship Closeness Scale, the lexical (but not facial array) measure of state anger is no longer significantly higher in the Sibling condition relative to the Acquaintance condition. The reason for this discrepancy between the lexical and facial results with regard to controlling for closeness is unclear. However, in this instance as well as in the many results which are unchanged when controlling for closeness, greater observed effects of the scenario target manipulation on ratings of anger, aggression, and disgust in response to harm

MORAL EMOTIONS ARE RELATIVE

998 befalling siblings versus friends cannot be driven by greater feelings of closeness with siblings,
999 as participants reported greater closeness with their friends.

1000 ⁵ Participants may have associated the lexical item “grossed out / disgusted” more with
1001 pathogen disgust than with sexual or moral disgust, as the state disgust lexical item demonstrated
1002 significant positive correlations with trait pathogen disgust and trait sexual disgust, whereas the
1003 state disgust facial array only significantly correlated with trait moral disgust (see SOM Table
1004 S10).

1005 ⁶ These percentage ranges refer to the facial array measures to permit comparability
1006 across all three studies (the lexical measures of fear and disgust were only used in Study 3).
1007