


# Punish or Protect? How Close Relationships Shape Responses to Moral Violations

Personality and Social Psychology Bulletin  
2020, Vol. 46(5) 693–708  
© 2019 by the Society for Personality and Social Psychology, Inc  
Article reuse guidelines:  
sagepub.com/journals-permissions  
DOI: 10.1177/0146167219873485  
journals.sagepub.com/home/pspb  


Aaron C. Weidman<sup>1,\*</sup> , Walter J. Sowden<sup>1,\*</sup> ,  
Martha K. Berg<sup>1</sup>, and Ethan Kross<sup>1</sup>

## Abstract

People have fundamental tendencies to punish immoral actors *and* treat close others altruistically. What happens when these tendencies collide—do people punish or protect close others who behave immorally? Across 10 studies ( $N = 2,847$ ), we show that people consistently anticipate protecting close others who commit moral infractions, particularly highly severe acts of theft and sexual harassment. This tendency emerged regardless of gender, political orientation, moral foundations, and disgust sensitivity and was driven by concerns about self-interest, loyalty, and harm. We further find that people justify this tendency by planning to discipline close others on their own. We also identify a psychological mechanism that mitigates the tendency to protect close others who have committed severe (but not mild) moral infractions: self-distancing. These findings highlight the role that relational closeness plays in shaping people's responses to moral violations, underscoring the need to consider relational closeness in future moral psychology work.

## Keywords

moral psychology, close relationships, loyalty, harm, self-distancing

Received March 19, 2019; revision accepted July 21, 2019

On February 5, 2018, Larry Nassar was sentenced to up to 175 years in prison after pleading guilty to seven counts of sexual assault of minors during his time as a physician working for USA Gymnastics and Michigan State University. One perplexing aspect of this case is that several seemingly upstanding members of the community who worked closely with Nassar appear to have enabled his criminal behavior rather than report his crimes (e.g., Barr & Murphy, 2018).

What makes this and countless other situations like it fascinating is that it pits two seemingly axiomatic tendencies against one another. On one hand, people have a fundamental tendency to condemn and punish acts that cause harm to others as highly immoral (e.g., Graham, Haidt, & Nosek, 2009; Hofmann, Brandt, Wisneski, Rockenbach, & Skitka, 2018; Schein & Gray, 2018). On the other hand, people have a fundamental tendency to look out for the welfare of close others (e.g., Aron, Aron, Tudor, & Nelson, 1991; Hamilton, 1964; Trivers, 1971). When people evaluate close others' moral infractions—as was the case for Larry Nassar's colleagues—these tendencies come into conflict, raising a pivotal question: Does one punish the immoral act or protect the close other?

Despite the serious implications of these dilemmas, they have received scant attention in the moral psychology literature. Instead, over the past two decades, moral psychology

has seen a proliferation of theoretical models and empirical studies speaking to how people react to moral violations involving strangers (e.g., trolley problems; vignettes describing a stranger having sex with a chicken; Bauman, McGraw, Bartels, & Warren, 2014; Bloom, 2011; Graham et al., 2009; Greene, Sommerville, Nystrom, Darley, & Cohen, 2001; Haidt, 2001; Schein & Gray, 2018). As a result, we know little about how people react to moral violations committed by close others.

This article aims to improve our understanding of this phenomenon by examining the degree to which close others influence reactions to moral violations, how these reactions vary depending on the severity of the violation people consider, whether they are influenced by theoretically relevant

<sup>1</sup>University of Michigan, Ann Arbor, USA

\*A.C.W. and W.J.S. are shared first authors.

## Corresponding Authors:

Aaron C. Weidman, Department of Psychology, University of Michigan, 530 Church St., Ann Arbor, MI 48109, USA.  
Email: weidman@umich.edu

Walter J. Sowden, Department of Behavioral Health, Tripler Army Medical Center, 1 Jarrett White Road, Honolulu, HI 96859, USA.  
Email: wjsowden@gmail.com

individual differences, and whether they can be mitigated through intervention.

### Resolving a Fundamental Tension: Punish or Protect Close Others' Immoral Acts

People from across the political spectrum view harmful acts as immoral (Graham et al., 2009; Schein & Gray, 2018). Yet it is unclear how people will respond toward close others who commit immoral acts. On one hand, people may take punitive action against close others following moral infractions—even if this creates conflict in the relationship—because doing so helps enforce norms and therefore promotes societal cohesion (Hofmann et al., 2018).

On the other hand, when people witness close others commit immoral acts, their first reaction may be to protect—rather than punish—the perpetrator even if punishment could affect long-term behavioral change. People have an instinctual tendency to care for close relationship partners (e.g., Aron et al., 1991). From an evolutionary standpoint, kin selection also suggests that humans are inclined to behave favorably toward genetically related individuals (Hamilton, 1964; for example, Burnstein, Crandall, & Kitayama, 1994), and reciprocal altruism leads people to act favorably toward close others even if they do not share genetic relatedness (Trivers, 1971; for example, Bernhard, Fischbacher, & Fehr, 2006). In line with these considerations, Waytz, Dungan, and Young (2013) provided initial evidence that people are less likely to report an immoral act to a third party (i.e., blowing the whistle) when considering acts of close others (i.e., family members) than distant others (i.e., acquaintances). Similarly, Hofmann and colleagues (2018) found that people reported less desire to punish immoral acts committed by perpetrators with whom they reported greater feelings of closeness.

The above considerations and initial empirical findings present a puzzle: Are people more likely to punish or protect close others whom they witness committing immoral acts? The first goal of this research was to examine the strength of people's tendency to protect, rather than punish, close others in the face of moral violations. To prioritize generalizability, we examined responses to moral violations of high and low severity and those that came from multiple moral domains: theft and sexual harassment. We further examined several potential mechanisms that could underlie the tendency to protect close others, including the extent to which individuals consider (a) their own self-interest, (b) loyalty, and (c) harm to the perpetrator versus harm to others in society (not including the perpetrator). For example, people might feel that protecting a close other is both in their self-interest and is loyal, whereas the harm that could arise from punishment might be more salient when considering a close (vs. distant) others' immoral act. Finally, we tested whether individual differences in moral foundations, disgust sensitivity, political orientation,

or gender attenuate the tendency to protect close others who have behaved immorally.

### Regulating Our Tendency to Protect Close Others' Immoral Acts

It is typically impossible to change the nature of our close relationships—a brother who commits a moral crime is a brother no less. If people wish to counteract the tendency to protect (vs. punish) a close other in the face of an immoral act, they may instead have to change their perspective on the close relationship. We propose that reflecting on experiences from a psychologically distanced perspective, a process we refer to as *self-distancing*, is a tool people might use to accomplish this goal (for review, see Kross & Ayduk, 2017).

Self-distancing is one technique by which people can create psychological distance from the self. Self-distancing allows people to focus on their self as though they were focusing on another person. By temporarily stepping back from the self—and in turn increasing psychological distance from one's relationship with a close other (Lieberman & Trope, 2008)—people may be able to make close others seem more akin to distant acquaintances. If self-distancing makes close others seem like distant others, it may in turn make people's reactions to close others' immoral acts more akin to their reactions to distant others' immoral acts, thereby reducing people's tendency to protect—and increasing people's tendency to punish—close others following immoral acts.

Supporting the possibility that self-distancing could increase people's tendency to punish close others who have behaved immorally, prior work has indicated that cueing people to judge immoral acts committed by strangers at a greater psychological distance (i.e., in the distant future as opposed to the near-future) leads people to evaluate these behaviors more harshly (e.g., Agerström & Björklund, 2013; Eyal, Liberman, & Trope, 2008). Other relevant work has shown that people who adopt a high-level (vs. low-level) construal mind-set are more likely to dissent by speaking out against the norms of groups with which they strongly identify (Packer, Fujita, & Chasteen, 2014).

Our second goal in the present work was therefore to examine whether self-distancing can be used as a tool to counteract the tendency to protect (vs. punish) close others after witnessing these individuals commit immoral acts. Yet, if self-distancing is effective in this manner, it may work best when people react to highly severe immoral acts. This is because considering how to respond after witnessing a close other commit a highly severe moral infraction is likely to be highly distressing, and prior work has shown that the benefits of self-distancing increase with the emotional intensity of the situation that people think about (e.g., Kross & Ayduk, 2009; Penner et al., 2016; Résibois, Kuppens, Van Mechelen, Fossati, & Verduyn, 2018). In fact, Penner and colleagues (2016) recently concluded that “a certain level of emotional

distress may be needed for self-distancing to exert its beneficial effects; if not enough distress is experienced, there may be little room for self-distancing to operate” (p. 636).

## The Present Research

Our article is divided into three parts encompassing 10 experiments. In Part 1 (Studies 1a-1e), we tested whether people are in fact more likely to protect close (vs. distant) others following immoral acts. We explored this behavior in the face of high- and low-severity immoral acts from across two moral domains (i.e., theft and sexual harassment). We examined acts of sexual harassment in light of the ongoing *Me Too* movement (Kantor & Twohey, 2017; Ohlheiser, 2017). We further examined whether several theoretically relevant individual differences attenuate this tendency and explored the morally relevant mechanisms underlying this tendency. In Part 2 (Studies 2a, 2b, and 3), we examined multiple possibilities for how people justify protecting close others who have behaved immorally. In Part 3 (Studies 4a and 4b), we examined whether self-distancing can be used to attenuate this tendency.

Across studies, we used a paradigm in which participants were asked to consider common moral violations committed by people whom they actually knew—as opposed to considering rare moral violations committed by complete strangers—to answer a call for more ecologically valid designs in moral psychology (Bauman et al., 2014; Bloom, 2011; Bostyn, Sevenhant, & Roets, 2018). Specifically, participants were asked to imagine that they observed an immoral act committed by a close or distant other, that they were then confronted by a police officer, and that they then had to decide whether to respond truthfully (punishing the perpetrator) or to lie (protecting the perpetrator). To the extent that people wish to protect close others after witnessing those individuals commit immoral acts, they will be more likely to predict lying to the police officer when the perpetrator is a close (vs. distant) other.

To prioritize replicability and generalizability, we used repeated measures within-subject and between-subjects designs, replicated our key findings in samples of undergraduate students in the lab and Amazon MTurk workers, and performed internal meta-analyses for key effects. All data and analysis scripts are publicly available (<https://osf.io/3dteu>). In the online supplement, we report how we determined our sample size; all data exclusions; all ancillary analyses, manipulations, and measures; and minor procedural differences across studies.

## Part I

### Method

**Overview.** We present methods for Studies 1a to 1e sequentially followed by a meta-analytic result of the effect of close

others on moral judgments. In Part 1, we sought to answer the following question: Are people more likely to punish or protect close others whom they witness committing immoral acts? We next present select results from Studies 1c, 1d, and 1e independently to answer two exploratory questions: (a) Do individual differences attenuate the effect of close others on moral judgments? and (b) Which psychological mechanisms mediate the effect of close others on moral judgments?

### Study 1a

**Participants.** Three hundred twenty-four participants completed the study, including 103 students (55% women; age data were not collected due to experimenter error) and 221 MTurk workers (56% women,  $M_{\text{age}} = 37.5$ ,  $SD = 12.6$ ).

**Procedure.** Participants were presented with nine scenarios in which each level of severity (low, moderate, high) was paired with each level of relational closeness (distant, casual, close) in a fully crossed,  $3 \times 3$  within-subjects design (see online supplement for description of how we generated immoral behavior scenarios and how participants generated relationship partners). Participants were asked to imagine that they had witnessed one of the individuals they had nominated commit an immoral act (e.g., credit card fraud, shoplifting, illegally downloading music). Participants were then asked to imagine that a police officer approached them and asked whether they knew anything about the immoral act. Participants could respond “yes,” indicating an intention to tell the truth and acknowledge that they witnessed the perpetrator’s unethical behavior, or “no,” indicating an intention to lie and claim not to have witnessed the perpetrator’s unethical behavior. We administered a binary dependent measure because it mirrored the dilemma faced by participants in daily reactions to these types of moral violations—whether to respond truthfully or lie to the police officer—thereby increasing ecological validity.

### Study 1b

**Participants.** Two hundred forty-eight MTurk workers participated (49% women;  $M_{\text{age}} = 37.9$ ;  $SD = 11.4$ ).

**Procedure.** To provide a targeted replication of Study 1a, we employed a  $2$  (close vs. distant others)  $\times$   $2$  (immoral behavior severity: high vs. low) between-subjects design. Participants were presented with a moral dilemma which had been rated in Study 1a as either low-severity (i.e., the “illegal download” dilemma) or high-severity (i.e., the “credit card fraud” dilemma; see Table S1). Depending on condition assignment, the perpetrator of this immoral act was a randomly chosen close or distant relationship partner.

As some participants might feel conflicted between the two binary response options used in Study 1a, in Study 1b participants were asked to indicate their predicted response

to the police officer by reporting the likelihood that they would respond that they had witnessed the immoral behavior (1 = “very unlikely”; 6 = “very likely”). This variable was reverse-coded; higher scores indicated greater predicted likelihood of lying to the police officer ( $M = 4.53$ ,  $SD = 1.84$ ).

### Study 1c

**Participants.** Six hundred and three MTurk workers participated (53% women;  $M_{\text{age}} = 36.7$ ;  $SD = 11.4$ ).

**Procedure.** Study 1c examined whether the effects in Studies 1a and 1b generalized to another domain by employing a 2 (relational closeness: close vs. distant)  $\times$  2 (severity: high vs. low)  $\times$  2 (immoral domain: theft or sexual harassment) between-subjects design, in which each participant responded to one moral infraction. Participants assigned to consider harassment-related immoral acts (e.g., groping, staring) were asked to nominate close or distant others who were male, given that the harassment scenarios were written to describe male protagonists (see online supplement for description of how we generated harassment immoral behavior scenarios).

To examine in an exploratory fashion whether individual differences moderated the effect of close others on protecting perpetrators of immoral acts, participants completed the Moral Foundations Questionnaire (MFQ; Graham et al., 2011) and Three Domains of Disgust Scale (TDD; Tybur, Lieberman, & Griskevicius, 2009; see also Olatunji et al., 2012) in random order after completing the moral violation paradigm. The MFQ yields five subscales (harm, fairness, loyalty, authority, and purity), each of which captures the extent to which individuals consider one moral foundation when making moral judgments, as well as how harshly individuals evaluate immoral behaviors related to that same foundation ( $\alpha$ s = .66-.89 across subscales;  $M$ s = 3.29-4.65;  $SD$ s = .77-1.35). The TDD yields three subscales (pathogen, sexual, and moral disgust), each of which captures how much disgust a person feels when considering immoral acts related to these three domains ( $\alpha$ s = .86-.91 across subscales;  $M$ s = 3.48-4.71;  $SD$ s = 1.18-1.47). Finally, participants reported their political orientation (1 = “strongly liberal”; 7 = “strongly conservative”;  $M = 3.40$ ,  $SD = 1.72$ ).

### Study 1d

**Participants.** Two hundred sixty-one MTurk workers participated (49% women;  $M_{\text{age}} = 37.5$ ,  $SD = 12.1$ ).

**Procedure.** The procedure of Study 1d built on Studies 1a to 1c by providing an initial exploratory examination of the psychological mechanisms that may lead people to protect close others who have committed immoral acts, including (a) self-interest (Aron, Aron, & Smollan, 1992), (b) loyalty (Graham et al., 2011), (c) harm (Schein & Gray, 2018), and (d) anger and disgust (Rozin, Lowery, Imada, & Haidt, 1999). We used

a 2 (relational closeness: close vs. distant)  $\times$  2 (severity: high vs. low) between-subjects design as in Study 1b. Participants were presented with a moral dilemma and then answered questions regarding how these mechanistic considerations may affect their decision of how to respond to the police officer. We did not ask participants to decide how they would respond to the police officer after completing these questions because the goal of Study 1d was to test the effect of close others on our set of proposed mechanisms.

**Proposed mechanisms.** We assessed the four proposed mechanisms described above with eight separate questions—two each for self-interest, loyalty, harm, and anger/disgust—each of which is analyzed independently below. The questions for each mechanism were always presented together (e.g., self-interest to punish and self-interest to protect), but the four sets of questions were presented in random order. Each question was answered on a 7-point scale (1 = “not at all”; 7 = “very much”).

**Self-interest.** Two questions assessed participants’ perceived self-interest in punishing the perpetrator (i.e., “To what extent is it in your own self-interest to tell the officer that you *did see* [perpetrator] commit the act in question?”) and protecting the perpetrator (i.e., “To what extent is it in your own self-interest to tell the officer that you *did not see* [perpetrator] commit the act in question?”;  $M_{\text{Punish}} = 3.02$ ,  $SD = 2.19$ ;  $M_{\text{Protect}} = 4.05$ ,  $SD = 2.25$ ). We assumed that participants would interpret self-interest in punishing the perpetrator as stemming from a desire to prevent the perpetrator from re-offending, and we assumed that participants would interpret self-interest in protecting the perpetrator as stemming from a desire to preserve the relationship with the perpetrator. The two items were negatively correlated ( $r = -.41$ ).

**Loyalty.** Two questions assessed participants’ perceived disloyalty in punishing the perpetrator (i.e., “To what extent is it disloyal to tell the officer that you *did see* [perpetrator] commit the act in question?”) and protecting the perpetrator (i.e., “To what extent is it disloyal to tell the officer that you *did not see* [perpetrator] commit the act in question?”;  $M_{\text{Punish}} = 4.47$ ,  $SD = 2.21$ ;  $M_{\text{Protect}} = 2.81$ ,  $SD = 2.04$ ). When considering the option of responding truthfully and punishing the perpetrator, we assumed that participants would interpret this question as gauging disloyalty to the perpetrator himself or herself. In contrast, when considering the option of lying to the police officer to protect the perpetrator, we assumed that participants would interpret this question as gauging disloyalty to other people in society who might be harmed if the perpetrator roams free and re-offends. The two items were negatively correlated ( $r = -.36$ ).

**Harm.** Two questions assessed whether participants were considering harm that would come from punishing the perpetrator (i.e., “How much *harm* [i.e., *physical or emotional*

**Table 1.** Effect of Close Others on Protecting Perpetrators of Immoral Acts (Studies 1a-1c and 1e).

Study	Overall effect	Effect at high-severity	Effect at low-severity
Study 1a	$b = 1.43, CI = [1.23, 1.63], p < .001$	$b = 1.85, CI = [1.52, 2.18], p < .001$	$b = 1.70, CI = [1.31, 2.09], p < .001$
Study 1b	$b = 1.22, CI = [0.79, 1.65], p < .001$	$b = 2.00, CI = [1.37, 2.63], p < .001$	$b = 0.58, CI = [0.11, 1.05], p = .02$
Study 1c: Theft	$b = 1.20, CI = [0.82, 1.60], p < .001$	$b = 1.58, CI = [1.03, 2.13], p < .001$	$b = 0.79, CI = [0.05, 1.53], p < .01$
Study 1c: Sexual harassment	$b = 1.01, CI = [0.64, 1.38], p < .001$	$b = 1.12, CI = [0.59, 1.65], p < .001$	$b = 0.83, CI = [0.34, 1.32], p < .001$
Study 1e	$b = 1.62, CI = [1.19, 2.05], p < .001$	$b = 2.00, CI = [1.41, 2.59], p < .001$	$b = 1.38, CI = [0.79, 1.97], p < .001$
Meta-analytic effect	$b = 1.34, CI = [1.20, 1.48], p < .001$	$b = 1.72, CI = [1.51, 1.94], p < .001$	$b = 1.10, CI = [0.89, 1.31], p < .001$

Note. Positive betas indicate that people were more likely to predict lying to the police officer to protect close (vs. distant) others who committed immoral acts. Effects are unstandardized betas on a 6-point scale. We also found a meta-analytic, two-way interaction between relational closeness and severity ( $b = 0.30, CI = [0.08, 0.52], p = .007$ ). CI = confidence interval.

suffering] would come to other people (including [perpetrator]) if you tell the officer that you *did see* [perpetrator] commit the act in question?”) or protecting the perpetrator (i.e., “How much *harm* [i.e., *physical or emotional suffering*] would come to other people (including [perpetrator]) if you tell the officer that you *did not see* [perpetrator] commit the act in question?”;  $M_{\text{Punish}} = 5.12, SD = 1.54; M_{\text{Protect}} = 3.24, SD = 2.02$ ). We assumed that participants who thought that harm would come from punishing the perpetrator were considering harm to the perpetrator, whereas participants who thought that harm would come from protecting the perpetrator were considering harm to others in society (we revised the wording of this question in Study 1e to avoid relying on this assumption; see online supplement ). These two items were relatively orthogonal ( $r = .02$ ).

**Anger and disgust.** Participants reported their feelings of anger and disgust at witnessing the immoral act using four-item scales (Harmon-Jones, Bastian, & Harmon-Jones, 2016;  $\alpha_s = .96$  and  $.92$  for anger and disgust, respectively;  $M_s = 3.41$  and  $2.76$ ;  $SD_s = 2.03$ - $1.77$ ).

### Study 1e

**Participants.** Two hundred sixty-one MTurk workers participated (60% women;  $M_{\text{age}} = 36.9, SD = 10.9$ ).

**Procedure.** Study 1e was identical to Study 1d in that we used a 2 (relational closeness: close vs. distant)  $\times$  2 (severity: high vs. low) between-subjects design, except that we also assessed our primary dependent measure from Studies 1b and 1c. Likelihood of responding truthfully to the police officer was again reverse-coded, such that higher scores indicated greater likelihood of lying to the police officer to protect the perpetrator ( $M = 4.01, SD = 1.95$ ).

### Results

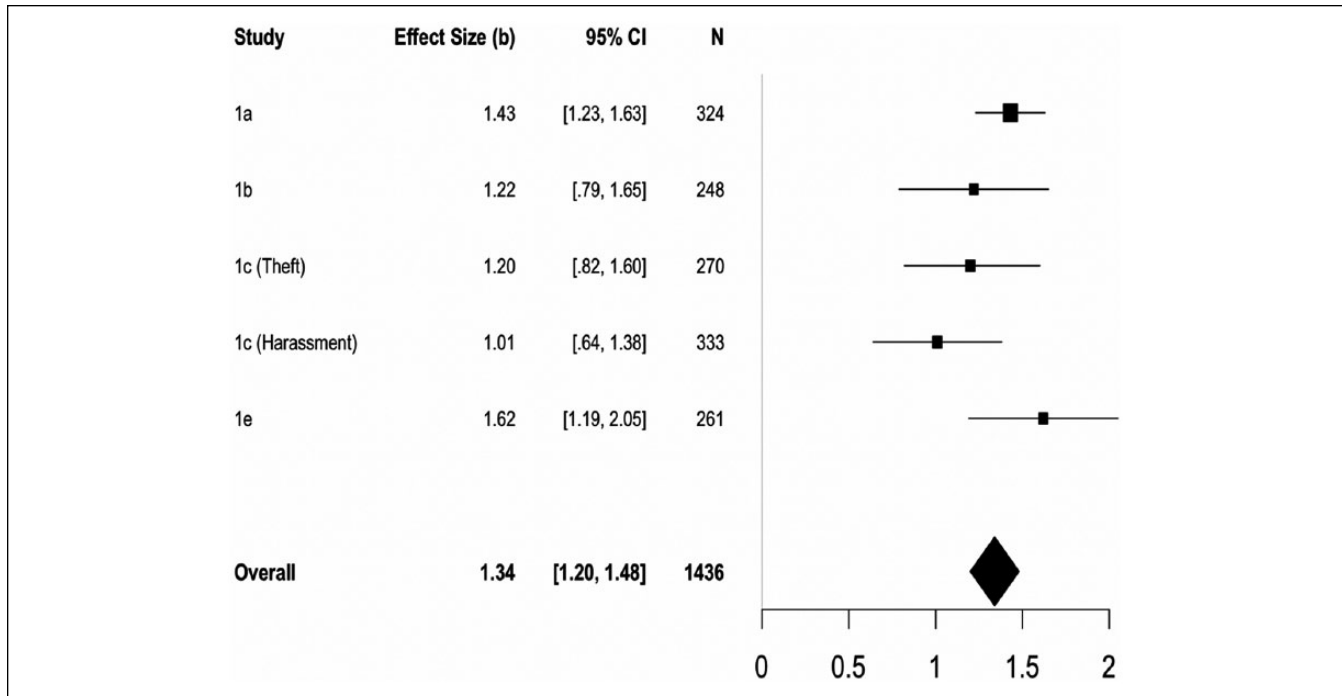
**Are people likely to protect close others following immoral acts? (Studies 1a-1c, 1e)?** Following Lipsey and Wilson (2001),

we conducted an internal meta-analysis across Studies 1a to 1c and 1e to arrive at a precise estimate of the effect size and to simplify presentation (participants did not report their response to the police officer in Study 1d; see Table 1 and Figure 1; see online supplement for full detail).

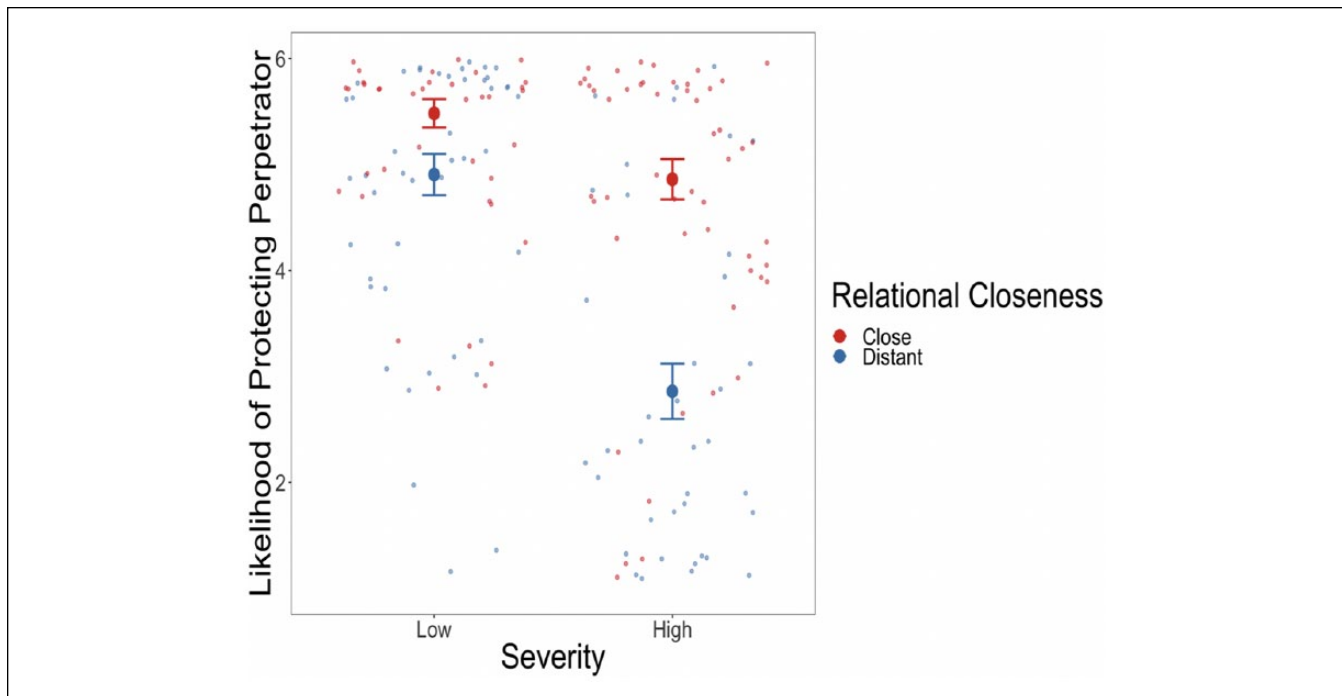
This internal meta-analysis revealed a strong overall effect: People were more likely to predict lying to a police officer to protect a close other who committed an immoral act compared with a distant other ( $b = 1.34$ , confidence interval [CI] = [1.20, 1.48],  $p < .001$ ); note that this effect corresponds to a 1.34-point difference on the 6-point scale used as our dependent measure. Importantly, this effect was statistically significant and of similar magnitude across studies and across moral violations that involved both theft and sexual harassment (Table 1). These findings suggest that people have a strong tendency to protect close others from ramifications of their immoral acts.

We also found that the effect of close others on lying to a police officer was stronger for high-severity violations ( $b = 1.72, CI = [1.51, 1.94], p < .001$ ) than low-severity violations ( $b = 1.10, CI = [0.89, 1.31], p < .001$ ; meta-analytic two-way interaction between closeness and severity:  $b = 0.30, CI = [0.08, 0.52], p = .007$ ). These findings suggest that people were most likely to predict protecting close others who committed immoral acts when those acts involved heinous acts of burglary, blackmail, groping, and proposing unwanted sex. Figure 2 shows these findings using data from Study 1b, which is representative of the pattern we observed across studies.

**Do individual differences moderate the influence of close relationships on protecting perpetrators of immoral acts (Study 1c)?** We next examined the interaction between each unique individual difference variable and relational closeness on predicted likelihood of lying to the police officer to protect the perpetrator, separately for theft and sexual harassment scenarios. We then examined the simple effect of close relationships on predicted lying to the police officer at both high (+1  $SD$ ) and low (−1  $SD$ ) levels of each individual difference variable (following Cohen, Cohen, West, & Aiken, 1991). This involved



**Figure 1.** Meta-analytic effect of close others on protecting perpetrators of immoral acts (Studies 1a-1c and 1e).  
 Note. Effects are unstandardized betas on a 6-point scale. Positive betas indicate a greater tendency to predict protecting close (vs. distant) others. Size of the square for each individual effect is inversely related to the width of the 95% confidence interval. Width of the diamond denoting overall effect represents the upper and lower bounds of the 95% confidence interval for this effect. Midpoint of the diamond centers on the overall effect size. Participants did not report the primary dependent measure in Study 1d.



**Figure 2.** Effects of close others on protecting perpetrators of immoral acts (Study 1b).  
 Note. Large dots represent means, and error bars around large dots represent  $\pm 1$  standard error. Small dots represent each individual data points.

computing 20 separate regressions, two for each of the five moral foundations subscales, three disgust sensitivity subscales, political orientation, and gender (see Table S4 for all statistics discussed in this section). Given the exploratory nature of these analyses, we used a Bonferroni correction to adjust our alpha level to .0025 (.05/20).

When adopting our Bonferroni-adjusted criterion for significance, we found zero instances in which an individual difference variable statistically moderated the effect of close others on predicted lying to the police officer. Bayesian analyses largely substantiated these null findings (see online supplement). Furthermore, in no case did the effect of relational closeness on likelihood of punishment become non-significant. The smallest effect size we observed emerged for individuals high on the purity foundation in response to sexual harassment scenarios ( $b = 0.59$ ,  $CI = [0.04, 1.14]$ ,  $p = .04$ ), yet this effect still indicates more than a half-point increase on our 6-point scale. We were also particularly struck by the fact that the effect of relational closeness remained strong and significant among individuals who reported *low* scores on the loyalty moral foundation ( $b = 0.75$ ,  $CI = [0.16, 1.34]$ ,  $p = .01$ ), even though *high* loyalty is a factor that should in theory drive the decision to protect close others who have behaved immorally.

Together, these findings indicate that both liberals and conservatives, men and women, and individuals with high and low levels of each domain of disgust sensitivity and each moral foundation—even people who reported that considerations of loyalty do not typically shape their moral concerns—were significantly more likely to predict that they would lie to a police officer to protect the perpetrator when considering close (vs. distant) others' immoral acts, across theft and harassment scenarios.

**Mediating mechanisms (Studies 1d and 1e).** We next examined the psychological processes that lead people to protect close others in these contexts. Below we report the findings from Study 1d (the findings for Study 1e were identical and are reported in the online supplement; see Tables S6 and S7).

**Self-interest.** People reported that they had more self-interest to *protect* the perpetrator when this individual was a close other compared with a distant other ( $M_{Close} = 4.64$ ,  $SD = 2.13$ ;  $M_{Distant} = 3.50$ ,  $SD = 2.22$ ;  $b = 1.14$ ,  $CI = [0.61, 1.67]$ ,  $p < .001$ ). This effect did not significantly differ across levels of severity (two-way interaction:  $b = 0.02$ ,  $CI = [-1.02, 1.08]$ ,  $p = .98$ ).

In contrast, people reported that they had more self-interest to *punish* the perpetrator when this individual was a distant other compared with a close other ( $M_{Close} = 2.76$ ,  $SD = 2.03$ ;  $M_{Distant} = 3.25$ ,  $SD = 2.31$ ;  $b = 0.49$ ,  $CI = [-0.04, 1.02]$ ,  $p = .07$ ). This effect did not significantly differ across levels of severity (two-way interaction:  $b = 0.57$ ,  $CI = [-0.47, 1.61]$ ,  $p = .28$ ).

These results suggest that people perceive protecting a close other, but punishing a distant other, as in their self-interest.

**Loyalty.** People reported that it was more disloyal to the perpetrator to *punish* the perpetrator when this individual was a close other compared with a distant other ( $M_{Close} = 5.30$ ,  $SD = 1.96$ ;  $M_{Distant} = 3.68$ ,  $SD = 2.16$ ;  $b = 1.62$ ,  $CI = [1.11, 2.13]$ ,  $p < .001$ ), although this effect was stronger when participants considered high (vs. low) severity immoral acts (high-severity:  $b = 2.29$ ,  $CI = [1.62, 2.96]$ ,  $p < .001$ ; low-severity:  $b = 1.12$ ,  $CI = [.41, 1.83]$ ,  $p < .01$ ; two-way interaction:  $b = 1.17$ ,  $CI = [0.19, 1.15]$ ,  $p = .02$ ).

In contrast, people reported that it was more disloyal to others in society to *protect* the perpetrator when this individual was a distant other compared with a close other ( $M_{Close} = 2.50$ ,  $SD = 1.85$ ;  $M_{Distant} = 3.10$ ,  $SD = 2.17$ ;  $b = 0.59$ ,  $CI = [0.10, 1.08]$ ,  $p = .02$ ). This effect did not differ across levels of severity (two-way interaction:  $b = 0.001$ ,  $CI = [-0.98, 0.98]$ ,  $p = .998$ ).

These results suggest that people perceive punishing a close other as a show of disloyalty to this individual, whereas they perceive protecting a distant other as disloyal to others in society at large.

**Harm.** People reported that more harm would come to the perpetrator from their decision to *punish* the perpetrator if he or she were a close other compared with a distant other ( $M_{Close} = 5.50$ ,  $SD = 1.48$ ;  $M_{Distant} = 4.76$ ,  $SD = 1.52$ ;  $b = 0.73$ ,  $CI = [0.36, 1.10]$ ,  $p < .001$ ). This effect did not significantly differ across levels of severity ( $b = 0.45$ ,  $CI = [-0.28, 1.18]$ ,  $p = .23$ ).

In contrast, people reported that more harm would come to others in society from their decision to *protect* the perpetrator if he or she were a distant other compared with a close other ( $M_{Close} = 2.94$ ,  $SD = 1.79$ ;  $M_{Distant} = 3.51$ ,  $SD = 2.18$ ;  $b = 0.57$ ,  $CI = [0.08, 1.06]$ ,  $p = .02$ ). This effect also did not significantly differ across levels of severity ( $b = 0.63$ ,  $CI = [0.13, 1.39]$ ,  $p = .11$ ).

These results suggest that harm to the perpetrator is a more salient consideration when people consider close others, whereas harm to other members of society is a more salient consideration when people consider distant others.

**Anger and disgust.** Close relationships did not affect participants' experience of anger ( $M_{Close} = 3.51$ ,  $SD = 2.03$ ;  $M_{Distant} = 3.32$ ,  $SD = 2.03$ ;  $b = 0.20$ ,  $CI = [0.29, 0.69]$ ,  $p = .44$ ) or disgust ( $M_{Close} = 2.73$ ,  $SD = 1.71$ ;  $M_{Distant} = 2.79$ ,  $SD = 1.83$ ;  $b = 0.06$ ,  $CI = [-0.37, 0.49]$ ,  $p = .78$ ), and neither of these effects differed across levels of severity (anger:  $b = 0.41$ ,  $CI = [-0.49, 1.31]$ ,  $p = .37$ ; disgust:  $b = 0.22$ ,  $CI = [-0.54, 0.98]$ ,  $p = .59$ ). These results suggest that anger and disgust are not affected by whether the perpetrator of the immoral act is a close or distant other.

**Full mediation model (Study 1e).** We next examined whether the link between close others and protecting the perpetrator of an immoral act was mediated by participants' concerns about self-interest, loyalty, and harm. Prior to conducting formal mediation analyses of the effect of close others on

protecting the perpetrator through the proposed mediating mechanisms, we examined whether the six mechanistic items capturing self-interest, loyalty, and harm grouped together into meaningful higher-order constructs. We performed an exploratory factor analysis on these items, using the *psych* package in R (Revelle, 2017). Following the results of a parallel analysis, we extracted three correlated factors using oblimin rotation (see Table S8).

Factor 1 appeared to represent *self-interest*, with a high loading for the item capturing self-interest to protect the perpetrator. As noted above, we interpreted this form of self-interest as stemming from concern over preserving the relationship with the perpetrator. Factor 2 appeared to capture *concern for harm to others in society*. This factor showed high loadings for the items capturing disloyalty in protecting the perpetrator (which, as noted above, can be viewed as disloyalty toward others in society who might be harmed if the perpetrator commits another immoral act), consideration of harm to others in society, and self-interest in punishing the perpetrator (which, as noted above, can be interpreted as self-interest in preventing the perpetrator from re-offending). Finally, Factor 3 appeared to capture *concern for harm to the perpetrator*, with high loadings on the items capturing consideration of harm to the perpetrator and disloyalty to the perpetrator that would come from punishing the immoral act.

Using scores on these three factors as mediators, we formally tested a mediational model through path analysis, using the *lavaan* package in R (Rosseel, 2012). In this model, relational closeness (randomly assigned) was treated as an independent variable, scores on self-interest, concern for society, and concern for the perpetrator were treated as mediators, and predicted likelihood of lying to the police officer was treated as the dependent variable. The model involved the following two steps: (a) the three mediators were regressed on relational closeness, and (b) predicted likelihood of lying to a police officer was regressed on the three mediators and relational closeness. The three mediators were allowed to correlate. Specifying our model in this manner yields simultaneous and independent statistical tests of the direct effect of close others on predicted lying to the police officer, as well as each indirect, mediational effect (e.g., the effect of close others on predicted lying to the police officer through self-interest).

In this model, close others had a positive, indirect effect on predicted likelihood of lying to the police officer through *increased* self-interest ( $b = 0.28$ ,  $CI = [0.12, 0.44]$ ,  $p < .01$ ), *increased* concern for harm to the perpetrator ( $b = 0.17$ ,  $CI = [0.03, 0.31]$ ,  $p < .01$ ), and *decreased* concern for harm to others in society ( $b = 0.62$ ,  $CI = [0.37, 0.87]$ ,  $p < .001$ ; see Figure 3). Close others also had a direct effect on predicted likelihood of protecting the perpetrator ( $b = 0.55$ ,  $CI = [0.18, 0.92]$ ,  $p < .01$ ), indicating that the aforementioned mechanistic factors only partially mediated this effect.

In sum, mediational analyses across Studies 1d and 1e suggested that people's tendency to protect close others after witnessing immoral acts is driven by perceived self-interest

and loyalty in protecting close others, as well as a tendency to try to minimize harm to close others who commit immoral acts even if doing so brings harm to society at large. In contrast, we did not find evidence that anger and disgust mediated the effect of close others on protecting the perpetrator of immoral acts.

## Part 2

We have documented the pervasive tendency for people to predict that they will protect close others from legal ramifications after witnessing these individuals commit immoral acts, particularly highly severe immoral acts of theft and sexual harassment. Yet overlooking immoral acts—even for the sake of preserving a cherished relationship—could create dissonance between one's sense of moral rectitude and the fact that one has failed to firmly punish a moral violation (Aronson, 1969). In Part 2, we sought to answer the following question: How do people justify failing to take punitive action in the face of moral violations? We considered the following two competing hypotheses in Studies 2 and 3.

On one hand, people may protect close others who commit immoral acts because close others shape moral judgments themselves. For example, if one's brother commits burglary, this act may be perceived as less immoral than if the act is committed by the mailman. As a result, protecting one's brother may seem justifiable, whereas turning one's brother in to law enforcement may seem like an overreaction.

On the other hand, people may judge close others' immoral acts as equivalently immoral but may choose to engage in a more lenient form of punishment that is aimed at making oneself feel morally justified while also preserving the relationship (e.g., Bandura, 2016). For example, one may consider a brother's burglary to be immoral but may wish to respond by talking to the brother about the seriousness of the act—which serves as a mild form of punishment for the act—rather than subjecting him to legal fallout with the police.

## Method

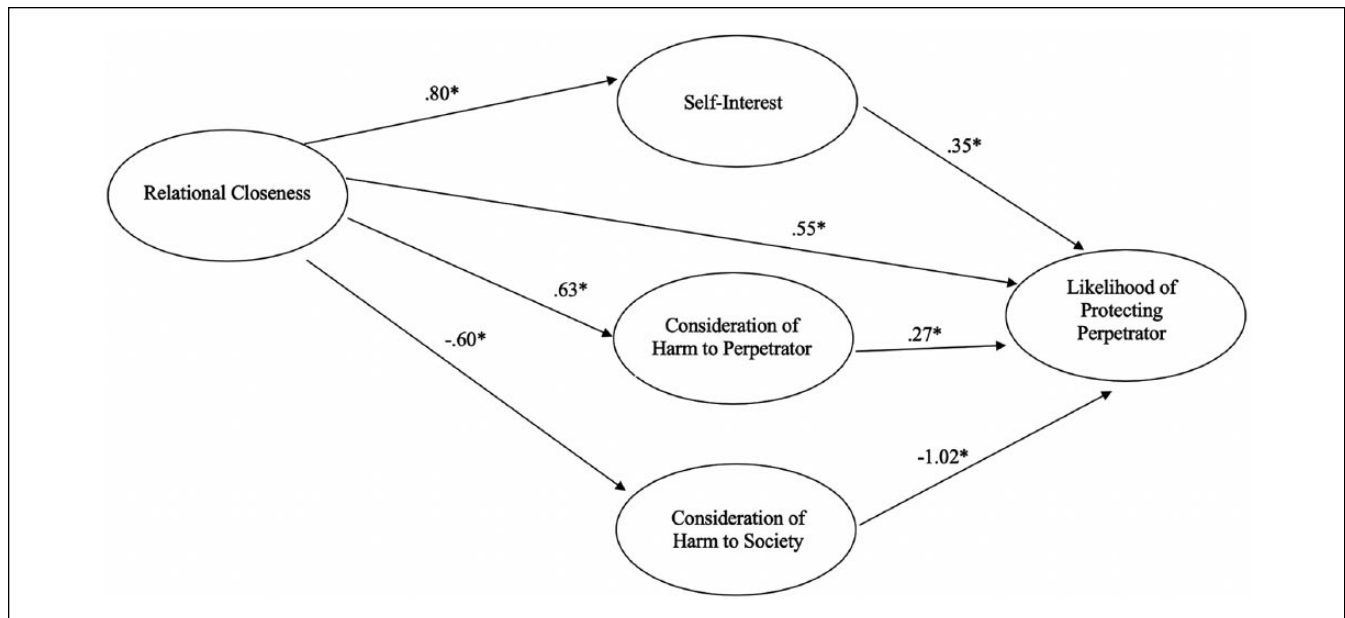
**Overview.** We present the methods for Studies 2a, 2b, and 3 sequentially. We next present results speaking to (a) whether close others shape moral judgments themselves or (b) whether people engage in more lenient forms of punishment after witnessing close others' immoral acts.

### Studies 2a and 2b

**Participants.** Four hundred and four adults participated through Prolific Academic (Study 2a:  $N = 206$ ; 64% women,  $M_{\text{age}} = 37.32$ ;  $SD = 10.73$ ; Study 2b:  $N = 198$ ; 67% women,  $M_{\text{age}} = 36.03$ ;  $SD = 12.33$ ).

**Procedure.** Using a 2 (between-subjects: close vs. distant others)  $\times$  2 (within-subjects, immoral behavior severity: high vs.





**Figure 3.** Meditational pathways of the effect of close others on protecting the perpetrator of immoral acts (Study 1e).

Note.  $N = 261$ . Path weights are unstandardized regression coefficients. Relational closeness is coded such that 0 = distant others and 1 = close others. For clarity, correlations among mediators are omitted from diagram (self-interest and consideration of harm to perpetrator:  $r = .15$ ; self-interest and consideration of harm to society:  $r = -.24$ ; consideration of harm to perpetrator and harm to society:  $r = -.29$ ). Confidence intervals and exact  $p$  values for all direct and indirect effects are presented in the main text.

\* $p < .05$ .

low) design similar to Studies 1b to 1e, participants were asked to consider two immoral acts. Depending on condition assignment, participants nominated two close or distant others using the same instructions as Studies 1b to 1e. Participants in Studies 2a and 2b considered the theft-related and harassment-related immoral acts, respectively, that were used in Study 1c (see Tables S2 and S3). Participants in Study 2b were therefore asked to nominate distant or close others who were male, as in Study 1c.

After considering each immoral act, participants made moral judgments using two items: (a) "How *wrong* is the act [perpetrator] committed?" (1 = "not at all wrong"; 7 = "very wrong"), and (b) "How *immoral* is the act [perpetrator] committed?" (1 = "not at all immoral"; 7 = "very immoral"). The name of one close or distant other who participants had nominated was piped in to the appropriate point in each question. These two items were averaged to form a composite of perceived moral wrongness ( $r = .89$ ). For purposes of catching inattentive respondents, participants were asked to write a few sentences describing how they arrived at their decision. Finally, participants completed the MFQ.

### Study 3

**Participants.** Ninety-nine adults participated through Prolific Academic (63% women,  $M_{\text{age}} = 36.57$ ;  $SD = 11.30$ ).

**Procedure.** Using the same paradigm as Study 2a, participants were asked to consider one theft-related immoral act in

a 2 (close vs. distant other)  $\times$  2 (high- vs. low-severity) between-subjects design. Vignettes differed from those used in Study 2a in one way: Participants were told that a police officer was also in the vicinity and appeared to have witnessed the act. We added this detail to mimic the scenario participants faced in Study 1, in which they were confronted by a police officer who asked whether they had witnessed the act. Unlike Study 1, however, in Study 3 participants were *not* told that the police officer approached them or asked about the act, to allow participants to voluntarily indicate whether they would involve the officer in punishing the perpetrator. After considering the immoral act, participants were asked to report how they would respond in this situation, using the following open-ended prompt: "What action would you take in response to witnessing [perpetrator]'s actions? Please consider actions right in the moment and down the line in your relationship with [perpetrator]."

Free responses were coded by two research assistants blind to our hypotheses for five themes, based in part on some exploratory coding conducted in Studies 1b and 4b, described in the online supplement: (a) confrontation (i.e., an intention to directly talk to the perpetrator about the immoral act), (b) reporting (i.e., an intention to report the perpetrator to either legal authorities or close others such as friends and family), (c) avoidance (i.e., an intention to create social distance between the perpetrator and oneself or one's social group), (d) support (i.e., an intention to help the perpetrator avoid possible negative consequences of the immoral act), and (e) nothing (i.e., an intention to take no

action; 0 = theme absent; 1 = theme present). The two original coders showed good agreement on each dimension (Gwet's ACs  $>.84$ ). We used Gwet's AC as an index of reliability instead of Cohen's kappa due to low base rates of observing each coding theme (Gwet, 2008; Spitznagel & Helzer, 1985). For cases in which these two coders disagreed ( $N = 45$ , or 5%), a third coder was asked to review the free responses to break the tie (see online supplement for full coding rubric).

## Results

*Do close others shape moral judgments themselves (Study 2)?* Aggregating across acts of theft and harassment and levels of severity, participants judged immoral acts committed by close others as equivalently immoral ( $M = 5.60$ ;  $SD = 1.09$ ) as those committed by distant others ( $M = 5.74$ ,  $SD = 0.93$ ;  $b = -0.14$ ,  $CI = [-0.33, 0.06]$ ,  $p = .18$ ). This effect was also small and non significant for both high-severity acts ( $b = -0.06$ ,  $CI = [-0.23, 0.11]$ ) and low-severity acts ( $b = -0.21$ ,  $CI = [-0.53, 0.10]$ ). Interestingly, for theft-related acts, participants did judge close (vs. distant) others' immoral acts as less severe ( $b = -0.29$ ,  $CI = [-0.56, -0.03]$ ,  $p = .03$ ), although this effect was driven by judgments of low-severity immoral acts (low-severity:  $b = -0.52$ ,  $CI = [-0.95, -0.09]$ ,  $p = .02$ ; high-severity:  $b = -0.06$ ,  $CI = [-0.29, 0.17]$ ,  $p = .59$ ). For harassment-related acts ( $b = 0.03$ ,  $CI = [-0.27, 0.32]$ ), moral judgments did not differ across close versus distant others (see Figures S3-S5).

Subsequent analyses revealed that none of the five moral foundation dimensions significantly moderated this effect. For example, even though feelings of loyalty toward the perpetrator led people to predict protecting close others in Study 1, moral judgments of close versus distant others were small and nonsignificant for individuals reporting both high ( $b = -0.18$ ,  $CI = [-0.46, 0.10]$ ,  $p = .21$ ) and low ( $b = -0.08$ ,  $CI = [-0.36, 0.20]$ ,  $p = .57$ ) scores on the loyalty moral foundation (two-way interaction:  $b = -0.01$ ,  $CI = [-0.05, 0.03]$ ,  $p = .63$ ; effects for high- and low-severity were contrast-coded).

*Do close others shape punitive responses to witnessing immoral acts (Study 3)?* When participants considered close others' immoral acts, the vast majority reported an intention to confront the perpetrator to discuss the act (79%,  $CI = [73%, 85%]$ ) and a small number of participants even voiced intention to support the perpetrator in the event of fallout from the act (13%,  $CI = [8%, 17%]$  for close others vs. only 2%,  $CI = [0%, 4%]$  for distant others; see Figure 4). In contrast, when participants considered distant others' immoral acts, only about half reported an intention to confront the perpetrator to discuss the act (55%,  $CI = [48%, 62%]$ ), which was significantly less than in the close others condition (79%,  $CI = [73%, 85%]$ ).

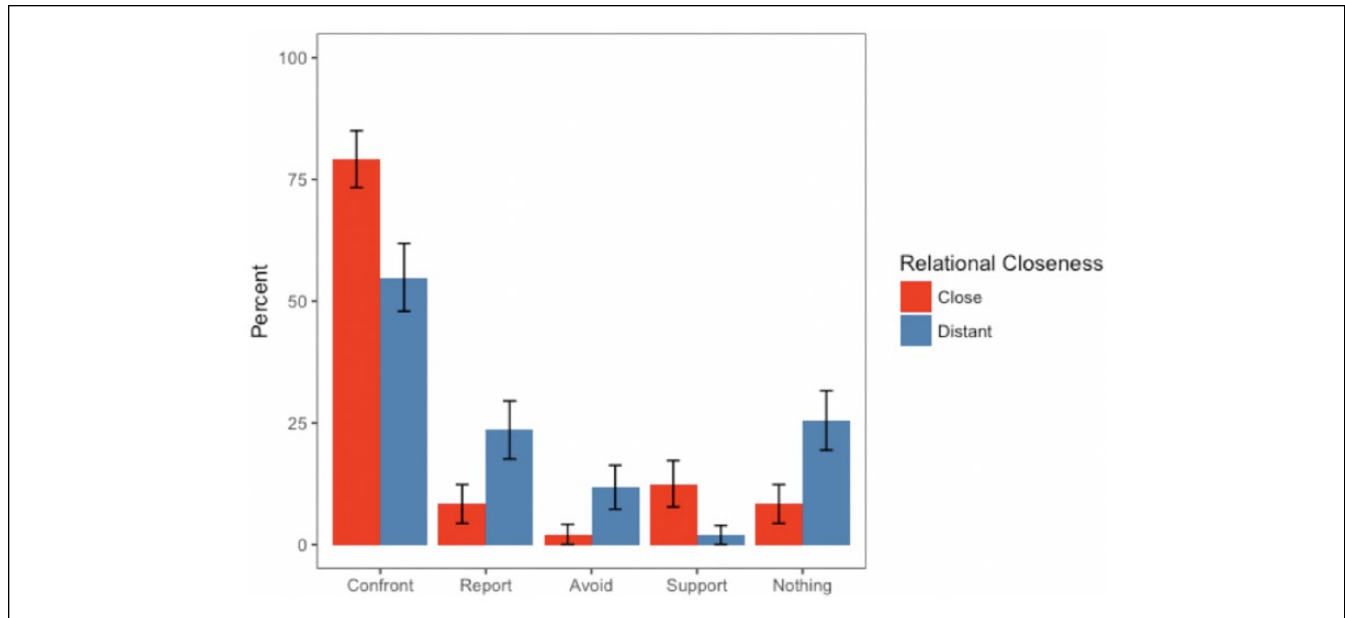
Among participants who considered distant others' immoral acts, a greater percentage reported an intention to take a harsher form of punitive action, such as reporting the perpetrator to law enforcement (24%,  $CI = [18%, 29%]$  vs. only 8%,  $CI = [4%, 12%]$ , for close others) or to actively avoid the perpetrator in future social interactions (12%,  $CI = [7%, 16%]$  vs. only 2%,  $CI = [0%, 4%]$ , for close others). Notably, a small proportion of participants who considered distant others' immoral acts reported an intention to take no action (25%,  $CI = [19%, 32%]$  vs. 8%,  $CI = [4%, 12%]$ , for close others).

*Summary.* Studies 2 and 3 suggest that participants' moral judgments themselves were relatively unaffected by close others: An act of burglary committed by one's brother appears to be viewed as equivalently immoral as an act of burglary committed by the mailman. Yet when people witness a close other commit an immoral act, they typically foresee engaging in a lenient form of punishment by confronting the perpetrator to discuss the act. In contrast, when people witness distant others commit immoral acts, these findings indicate that they are much more likely to take punitive action that will damage or sever the relationship, such as turning the perpetrator in to law enforcement or distancing themselves from the perpetrator.

## Part 3

Study 1 demonstrated a strong and pervasive tendency for people to protect close others who have committed immoral acts from legal ramifications. Studies 2 and 3 shed light on how people might justify overlooking immoral acts in this manner: Although close others' immoral acts are still viewed as immoral, people prefer to take more lenient recourse after witnessing these acts, with the goal of preserving the relationship. Although preserving close relationships is a very reasonable goal—and one toward which humans are naturally inclined (e.g., Aron et al., 1991; Hamilton, 1964; Trivers, 1971)—one can envision scenarios in which lying to a police officer to protect a close other could turn out very poorly (e.g., the participant who lied to the officer is subject to his or her own legal fallout; the perpetrator commits another immoral act). As a result, people may at times wish to lessen the tendency to protect close others who have acted immorally.

In Part 3, we therefore sought to answer the following question: Can self-distancing be used as a tool to counteract the tendency to protect (vs. punish) close others after witnessing these individuals commit immoral acts? As in Study 1, participants were randomly assigned to predict how they would respond to a police officer while adopting a self-immersed perspective (i.e., by using first-person pronouns) or a self-distanced perspective (i.e., by using non-first-person pronouns and one's name; following Kross et al., 2014).



**Figure 4.** Effect of close others on punitive tactics after witnessing immoral acts (Study 3).

Note. Percent corresponds to the percentage of participants in each relational closeness condition (close vs. distant) who wrote that they intended to engage in the specified punitive tactic.

As noted earlier, we predicted a priori that self-distancing may be particularly effective when people consider high-severity immoral acts (e.g., Kross & Ayduk, 2009; Penner et al., 2016; Résibois et al., 2018).

### Method

**Overview.** We present the methods for Studies 4a and 4b sequentially before presenting the results speaking to whether self-distancing mitigates the tendency to protect close others who have behaved immorally.

#### Study 4a

**Participants.** Four hundred and six participants completed the study, including 206 University of Michigan students (55% women) and 200 MTurk workers (57% women). Age information was not collected due to experimenter error.

**Procedure.** Using a similar paradigm as Study 1a, participants considered moral infractions in a 3 (within-subjects, relational closeness: close vs. casual vs. distant) by 3 (within-subjects, severity: low vs. medium vs. high) by 2 (between-subjects, perspective: self-immersed vs. self-distanced) design.

Following Kross et al. (2014), to manipulate self-distancing, participants were randomly assigned to think about their decision using either first-person, singular pronouns (a self-immersed perspective; for example, “What facts am I considering when making this decision?”) or third-person pronouns and their own name (a self-distanced perspective;

for example, “What facts is Walt considering when making this decision?”). After reading these instructions, participants completed the same binary dependent measure of how to respond to the police officer that was used in Study 1a.

#### Study 4b

**Participants.** Two hundred forty-nine University of Michigan students completed the study (59% women;  $M_{\text{age}} = 19.2$ ;  $SD = 1.11$ ).

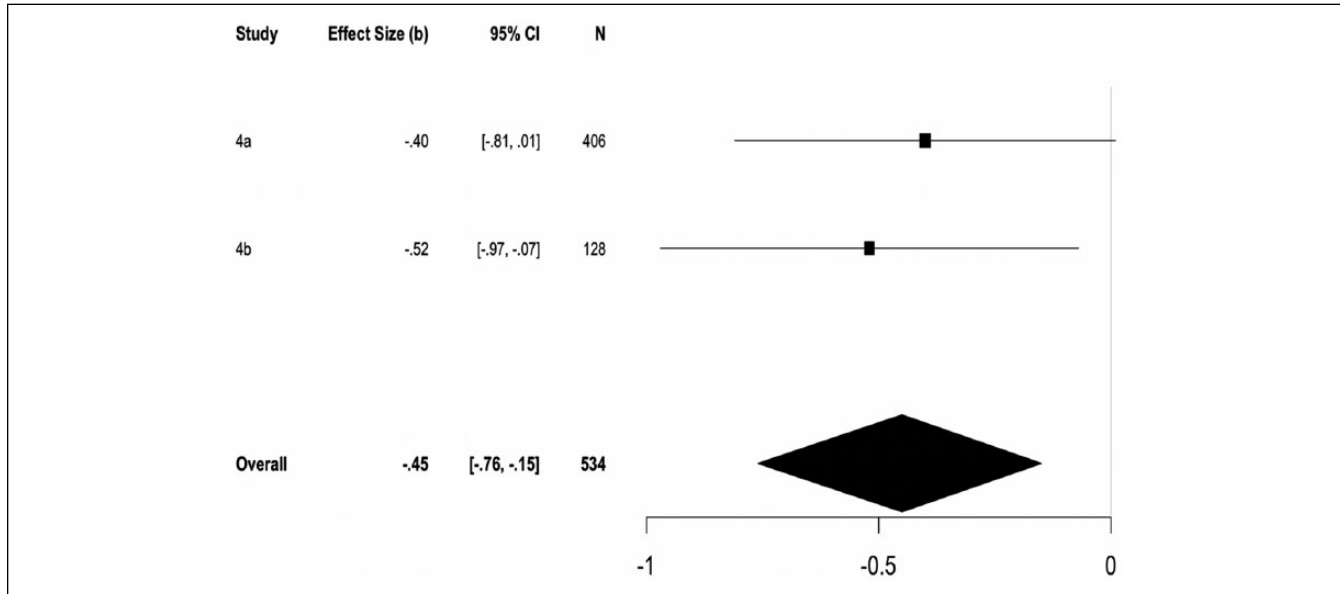
**Procedure.** Study 4b aimed to provide a targeted replication of the key findings in Study 2a using a  $2 \times 2$  between-subjects design in which participants responded to moral infractions of high or low severity while adopting a self-distanced or self-immersed perspective. Following the finding in Study 4a that the effect of self-distancing on protecting the perpetrator emerged for close others (but not distant others; see below), in Study 4b all immoral acts were committed by close others. In Study 4b, we also manipulated self-distancing via writing as in Study 1b, rather than silent introspection as in Study 4a, to test whether the effect of self-distancing emerges across modalities (see online supplement for a manipulation check).

Participants next completed the continuous dependent measure of how to respond to the police officer from Study 1b rather than the binary measure used in Study 4a, which was again reverse-coded such that higher scores indicated greater predicted likelihood of protecting a close other by lying to the police officer ( $M = 4.50$ ,  $SD = 1.57$ ).

**Table 2.** Internal Meta-Analysis of the Effect of Self-Distancing on Protecting Close Others (Studies 4a and 4b).

Study	Overall effect	Effect at high-severity	Effect at low-severity
Study 4a	$b = -0.32$ , CI = $[-1.01, 0.37]$ , $p = .37$	$b = -0.40$ , CI = $[-0.81, 0.01]$ , $p = .06$	$b = 0.11$ , CI = $[-2.58, 2.80]$ , $p = .93$
Study 4b	$b = -0.20$ , CI = $[-0.55, 0.15]$ , $p = .28$	$b = -0.52$ , CI = $[-0.97, -0.07]$ , $p = .02$	$b = 0.16$ , CI = $[-0.29, 0.61]$ , $p = .49$
Meta-Analytic Effect	$b = -0.23$ , CI = $[-0.55, 0.09]$ , $p = .16$	$b = -0.45$ , CI = $[-0.76, -0.15]$ , $p < .01$	$b = .16$ , CI = $[-0.29, 0.60]$ , $p = .48$

Note. Negative betas indicate that people were less likely to predict lying to the police officer to protect close others if they adopted a self-distanced (vs. self-immersed) perspective when considering moral infractions. Effects are unstandardized betas on a 6-point scale. We also found a meta-analytic, two-way interaction between self-distancing and severity at high relational closeness ( $b = .57$ , CI =  $[.20, .95]$ ,  $p = .003$ ). CI = confidence interval.

**Figure 5.** Effect of self-distancing on protecting close others following high-severity moral violations (Studies 4a and 4b).

Note. Effects are unstandardized betas on a 6-point scale. Negative betas indicate less likelihood of predicted lying to a police officer to protect close others. Size of the square for each individual effect is inversely related to the width of the 95% CI. Width of the diamond denoting overall effect represents the upper and lower bounds of the 95% CI for this effect. CI = confidence interval.

*Does self-distancing attenuate the effect of close others on protecting perpetrators of immoral acts (Studies 4a and 4b)?* We meta-analyzed the effect of self-distancing on protecting close others for both high- and low-severity violations across Studies 4a and 4b (Table 2 and Figure 5; see online supplement for full detail). This amounted to examining a two-way interaction between self-distancing and severity at high relational closeness. In line with this goal, in Study 4a we examined only responses when close others were nominated as perpetrators, and Study 4b was specifically designed so that participants nominated only close others as perpetrators.

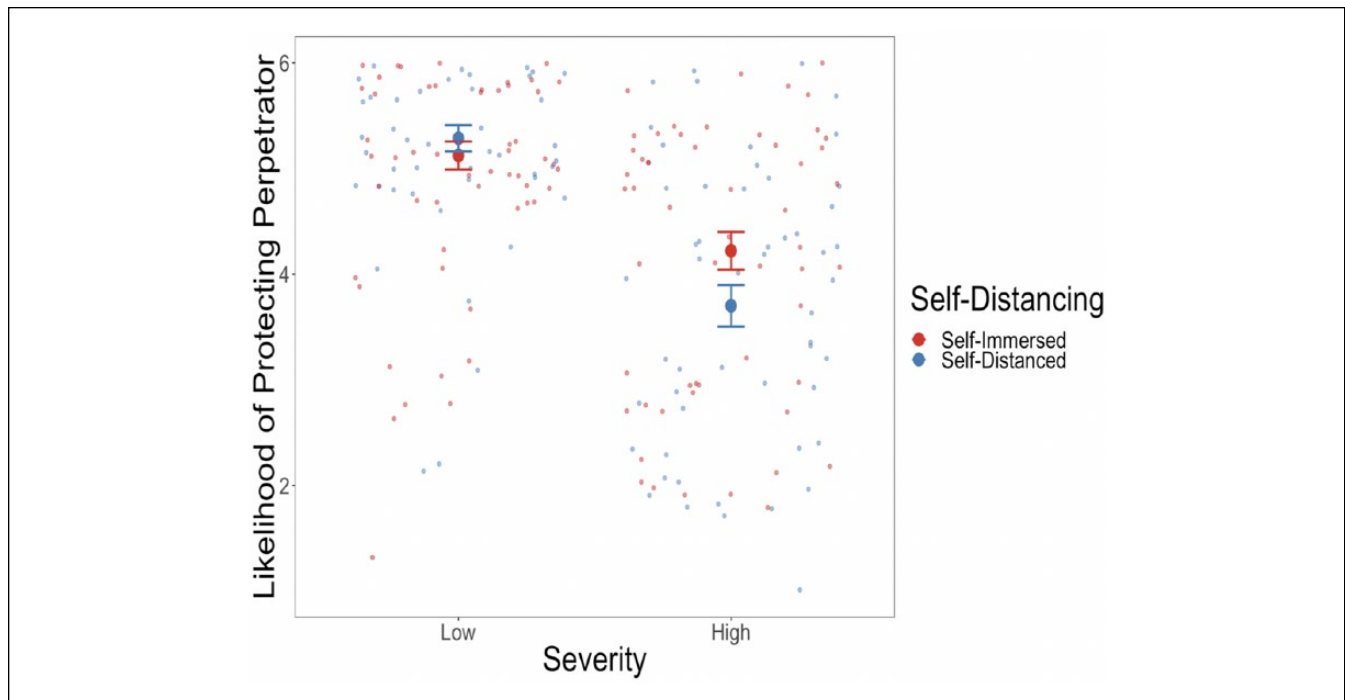
When people considered *high-severity* immoral acts of close others, self-distancing significantly reduced the tendency for people to predict lying to the police officer to protect those close others ( $b = -0.45$ , CI =  $[-0.76, -0.15]$ ,  $p < .01$ ). In contrast, when people considered *low-severity* immoral acts of close others, the effect of self-distancing was nonsignificant ( $b = 0.16$ , CI =  $[-0.29, 0.60]$ ,  $p = .48$ ; meta-analytic two-way interaction between self-distancing and

severity at high relational closeness:  $b = 0.57$ , CI =  $[0.20, 0.95]$ ,  $p = .003$ ). Figure 6 depicts each of these findings using data from Study 4b, which is also representative of Study 4a.

These results suggest that self-distancing is a context-bound tool that can help curb individuals' tendency to protect close others who have committed immoral acts in perhaps the most consequential moral scenarios: when those immoral acts are highly severe. At the same time, these findings also suggest that self-distancing is unlikely to sway people's decisions to protect close others in the face of low-severity immoral acts, indicating a critical boundary condition for the effect of self-distancing.

## General Discussion

It may seem natural to express outrage and disbelief when one learns that several of Larry Nassar's close colleagues failed to report his criminal behavior (Barr & Murphy, 2018).



**Figure 6.** Effect of self-distancing on protecting close others (Study 4b).

Note. Large dots represent means, and error bars around large dots represent  $\pm 1$  standard error. Small dots represent each individual data points.

We suspect that most readers would cringe at the notion that they themselves would flout the norms of society by helping to cover up such heinous acts. Yet, cringe or not, the current findings suggest that, when confronted with this type of moral dilemma, most people predict that they will behave *the same way* by protecting close others who have committed moral infractions.

We observed that the tendency to protect close others who have behaved immorally is, somewhat shockingly, most pronounced when people consider particularly heinous acts such as burglary, blackmail, groping, and proposing unwanted sex. This effect is large, corresponding to over a full point on a 6-point Likert-type scale (see Figure 1). Moreover, in exploratory analyses of potential individual difference moderators, it emerged in men and women, individuals from across the political spectrum, and people who place differential importance on the five moral foundations—including people who placed little importance on the loyalty foundation (Graham et al., 2009). We also used exploratory analyses to uncover the social-cognitive bases for this tendency: Protecting close others who have behaved immorally is seen as an act of self-interest (in that it preserves the relationship) as well as loyalty toward the perpetrator and is driven by a desire not to bring harm to a close other even at the cost of harm to society at large.

These findings also clarify the mechanisms through which people reconcile behaving loyally (by protecting close others who commit moral infractions) at the cost of behaving dishonestly while allowing an immoral actor to evade formal

punishment (by lying to a police officer). It does not appear that people view close others' moral infractions as less immoral: A brother's heinous crime is still a heinous crime. Instead, when people observe close others behaving immorally, we found through an exploratory linguistic coding analysis that they overwhelmingly intend to enact a lenient form of punishment by confronting the perpetrator to discuss the act. We suspect that doing so allows a person to simultaneously (a) maintain their self-image as a morally upstanding individual and (b) preserve and even enhance the close relationship, in line with the finding in Studies 1d and 1e that protecting close others from legal fallout is viewed as an act of self-interest. These tactics are also broadly consistent with prior work suggesting that people often justify their own immoral acts by focusing on positive consequences of the act or reaffirming their own moral standing (Bandura, 2016). In contrast, we found that when people observe distant others behaving immorally, they report greater intentions to subject these individuals to external, formal means of punishment, such as turning them in to law enforcement or subjecting them to social ostracization.

Finally, we observed that a brief self-distancing intervention reduced people's tendency to protect close others in the face of highly (but not moderately) severe immoral acts—the exact context in which this tendency is most pernicious. The context-bound effect of self-distancing on reactions to moral infractions emerged in two studies and in an internal meta-analysis—attesting to its robustness—and is consistent with an accumulating body of work indicating that the benefits of

self-distancing emerge in high-distress conditions (for discussion, see Kross & Ayduk, 2017)

## Theoretical Implications

These findings speak to the importance of considering close others when examining reactions to moral transgressions. The vast majority of prior work in moral psychology speaks to people's reactions to moral violations committed by unnamed strangers (e.g., Graham et al., 2009; Greene et al., 2001; Haidt, 2001; Schein & Gray, 2018). Yet, the present work suggests that reactions to close (vs. distant) others' immoral acts fundamentally differ: They are more lenient and relationship-focused, and they involve different weighting of social-cognitive considerations such as self-interest, loyalty, and harm. These findings suggest that if theoretical models of moral judgment and punishment continue to be based solely on strangers' violations, they will inaccurately predict how people will react when they are faced with the common occurrence of observing close others' immoral acts (see Bauman et al., 2014; Bloom, 2011; Bostyn et al., 2018, for similar arguments).

Our findings also have implications for research and theory on self-distancing and self-talk. To date, research in this domain has focused on how using one's own name and other non-first-person pronouns influences people's ability to regulate their emotions and reason wisely about social dilemmas (for review, see Kross & Ayduk, 2017). The current findings extend this research to the domain of moral judgments while adding to a growing body of research that suggests that self-distancing is particularly useful for helping people regulate their emotions in emotionally intense situations (e.g., Kross & Ayduk, 2009; Penner et al., 2016; Résibois et al., 2018). Future work is needed to explore the applied implications of subtle shifts in language for regulating our moral decision making.

## Future Directions

We assessed individuals' predicted responses (rather than actual responses) to observing close others' immoral behavior. We consistently observed that people predicted that they would respond untruthfully to a police officer to protect a close other who had behaved immorally, which amounts to people predicting that they would break the law by deliberately lying. If people were to mispredict their responses to these dilemmas, it is likely that they would underpredict the extent to which they would lie to the police officer to save face, because this behavior is socially undesirable (e.g., Paulhus, 1991). Nevertheless, future research should examine how our findings generalize to immoral behavior both inside and outside the laboratory.

Future work could also delve more deeply into the mechanisms driving the effect of relational closeness on moral punishment. For example, although we found that self-interest

underlies this effect, self-interest could stem from multiple motivations. One of these is a desire to preserve one's relationship with a close other—our proposed explanation. Yet self-interest could also stem from a desire to avoid reputational harm that could come from associating with a close other who is publicly known to have behaved immorally—that is, if we report a close other to a police officer, it could reflect poorly on our choice of friends.

Additional mechanisms, apart from the ones explored in this article, could also prove fruitful in explaining this effect. One example is perspective-taking: When we witness a close other behave immorally, we may feel that we know that person so well that we can predict with certainty that he or she will not re-offend. We also may feel so acquainted with a close other that we can imagine compelling (and therefore mitigating) reasons that he or she may have behaved immorally. This process of feeling like we can get inside the head of a close other who has behaved immorally may lead us to try and handle the dilemma privately (as observed in Study 3) rather than publicly, such as by reporting the incident to law enforcement. Future work should explore this possibility.

Finally, although we found that people were typically reluctant to report a close other's immoral behavior to a police officer, it would be fascinating to learn what would happen across a wider array of informants (e.g., parents, romantic partners, friends). For example, people may be more likely to report a close other's immoral behavior to a trusted confidant such as a parent or romantic partner, because that confidant could help reinforce more desirable behavior in the future. Reporting a teenager's petty attempt to illegally download music to a parent might be viewed as akin to privately reprimanding the teenager in a manner that will have positive remedial consequences—as opposed to involving law enforcement—which we observed in Study 3 is an approach most people take when they witness close others behaving immorally.

## Coda

These findings highlight the critical need to consider how close others shape reactions to moral infractions. They further suggest that doing so is essential for understanding the range of moral judgments, reactions, and punishments people undertake in everyday life.

## Authors' Note

Walter Sowden is now at Department of Behavioral Health, Tripler Army Medical Center, Honolulu, HI, USA.

## Author Contributions

W.J.S and E.K. developed the study concept and designed and conducted Studies 1a-1b and 4a-4b. A.C.W. and E.K. designed and conducted Studies 1c to 1e. A.C.W., M.K.B., and E.K. designed and conducted Studies 2 and 3. A.C.W. and M.K.B. analyzed the data. A.C.W. wrote the article, with assistance from E.K. Critical

revisions were provided by W.J.S., M.K.B., and E.K. All authors approved the article for submission.


### 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) received no financial support for the research, authorship, and/or publication of this article.

### ORCID iDs

Aaron C. Weidman  <https://orcid.org/0000-0002-9395-4051>

Walter J. Sowden  <https://orcid.org/0000-0001-7329-1222>

### Supplemental material

Supplemental material is available online with this article.

### References

- Agerström, J., & Björklund, F. (2013). Why people with an eye toward the future are more moral: The role of abstract thinking. *Basic and Applied Social Psychology, 35*, 373-381.
- 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*, 596-612.
- Aron, A., Aron, E. N., Tudor, M., & Nelson, G. (1991). Close relationships as including other in the self. *Journal of Personality and Social Psychology, 60*, 241-253.
- Aronson, E. (1969). The theory of cognitive dissonance: A current perspective. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 4, pp. 1-34). San Diego, CA: Academic Press.
- Bandura, A. (2016). *Moral disengagement: How people do harm and live with themselves*. New York, NY: Worth Publishers.
- Barr, J., & Murphy, D. (2018, January 16). Nasser surrounded by adults who enabled his predatory behavior. *ESPN*. Retrieved from [http://www.espn.com/espn/otl/story/\\_/id/22046031/michigan-state-university-doctor-larry-nassar-surrounded-enablers-abused-athletes-espn](http://www.espn.com/espn/otl/story/_/id/22046031/michigan-state-university-doctor-larry-nassar-surrounded-enablers-abused-athletes-espn)
- Bauman, C. W., McGraw, A. P., Bartels, D. M., & Warren, C. (2014). Revisiting external validity: Concerns about trolley problems and other sacrificial dilemmas in moral psychology. *Social and Personality Psychology Compass, 8*, 536-554.
- Bernhard, H., Fischbacher, U., & Fehr, E. (2006). Parochial altruism in humans. *Nature, 442*, 912-915.
- Bloom, P. (2011). Family, community, trolley problems and the crisis in moral psychology. *The Yale Review, 99*, 26-43.
- Bostyn, D. H., Sevenhant, S., & Roets, A. (2018). Of mice, men, and trolleys: Hypothetical judgments versus real-life behavior in trolley-style moral dilemmas. *Psychological Science, 29*, 1084-1093.
- Burnstein, E., Crandall, C., & Kitayama, S. (1994). Some Neo-Darwinian decision rules for altruism: Weighing cues of inclusive fitness as a function of the biological importance of the decision. *Journal of Personality and Social Psychology, 67*, 773-789.
- Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (1991). *Applied multiple regression/correlation analysis for the behavioral sciences*. Hillsdale, NJ: Lawrence Erlbaum.
- Eyal, T., Liberman, N., & Trope, Y. (2008). Judging near and distant virtue and vice. *Journal of Experimental Social Psychology, 44*, 1204-1209.
- Graham, J., Haidt, J., & Nosek, B. (2009). Liberals and conservatives relay on different sets of moral foundations. *Journal of Personality and Social Psychology, 96*, 1029-1046.
- Graham, J., Nosek, B. A., Haidt, J., Iyer, R., Koleva, S., & Ditto, P. H. (2011). Mapping the moral domain. *Journal of Personality and Social Psychology, 101*, 366-385.
- Greene, J. D., Sommerville, R. B., Nystrom, L. E., Darley, J. M., & Cohen, J. D. (2001). An fMRI investigation of emotional engagement in moral judgment. *Science, 293*, 2105-2108.
- Gwet, K. L. (2008). Computing inter-rater reliability and its variance in the presence of high agreement. *British Journal of Mathematical and Statistical Psychology, 61*, 29-48.
- Haidt, J. (2001). The emotional dog and its rational tail: A social intuitionist approach to moral judgment. *Psychological Review, 108*, 814-834.
- Hamilton, W. D. (1964). The genetical evolution of social behaviour. II. *Journal of Theoretical Biology, 7*, 17-52.
- Harmon-Jones, C., Bastian, B., & Harmon-Jones, E. (2016). The discrete emotions questionnaire: A new tool for measuring state self-reported emotions. *PLoS ONE, 11*, e0159915.
- Hofmann, W., Brandt, M. J., Wisneski, D. C., Rockenbach, B., & Skitka, L. J. (2018). Moral punishment in everyday life. *Personality and Social Psychology Bulletin, 44*, 1697-1711.
- Kantor, J., & Twohey, M. (2017, October 5). Harvey Weinstein paid off sexual harassment accusers for decades. *The New York Times*. Retrieved from <https://www.nytimes.com/2017/10/05/us/harvey-weinstein-harassment-allegations.html>
- Kross, E., & Ayduk, O. (2009). Boundary conditions and buffering effects: Does depressive symptomology moderate the effectiveness of self-distancing for facilitating adaptive emotional analysis? *Journal of Research in Personality, 43*, 923-927.
- Kross, E., & Ayduk, O. (2017). Self-distancing: Theory, research and current directions. *Advances in Experimental Social Psychology, 55*, 81-136.
- Kross, E., Bruehlman-Senecal, E., Park, J., Burson, A., Dougherty, A., Shablack, H., . . . Ayduk, O. (2014). Self-talk as a regulatory mechanism: How you do it matters. *Journal of Personality and Social Psychology, 106*, 304-324.
- Liberman, N., & Trope, Y. (2008). The psychology of transcending the here and now. *Science, 322*, 1201-1205.
- Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. Thousand Oaks, CA: SAGE.
- Ohlheiser, A. (2017, October 19). Meet the woman who coined "Me Too" 10 years ago—To help women of color. *The Washington Post*. Retrieved from <http://www.chicagotribune.com/lifestyles/ct-me-too-campaign-origins-20171019-story.html>
- Olatunji, B. O., Adams, T., Ciesielski, B., David, B., Sarawgi, S., & Broman-Fulks, J. (2012). The three domains of disgust scale: Factor structure, psychometric properties, and conceptual limitations. *Assessment, 19*, 205-225.
- Packer, D. J., Fujita, K., & Chasteen, A. L. (2014). The motivational dynamics of dissent decisions: A goal-conflict approach. *Social Psychological and Personality Science, 5*, 27-34.

- Paulhus, D. L. (1991). Measurement and control of response bias. In J. P. Robinson, P. R. Shaver, & L. S. Wrightsman (Eds.), *Measures of personality and social psychological attitudes* (pp. 17-59). San Diego, CA: Academic Press.
- Penner, L. A., Guevarra, D. A., Harper, F. W., Taub, J., Phipps, S., Albrecht, T. L., & Kross, E. (2016). Self-distancing buffers high trait anxious pediatric cancer caregivers against short-and longer-term distress. *Clinical Psychological Science, 4*, 629-640.
- Résibois, M., Kuppens, P., Van Mechelen, I., Fossati, P., & Verduyn, P. (2018). Depression severity moderates the relation between self-distancing and features of emotion unfolding. *Personality and Individual Differences, 123*, 119-124.
- Revelle, W. (2017). *psych: Procedures for personality and psychological research*. Evanston, IL: Northwestern University.
- Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software, 48*, 1-36.
- Rozin, P., Lowery, L., Imada, S., & Haidt, J. (1999). The CAD triad hypothesis: A mapping between three moral emotions (contempt, anger, and disgust) and three moral codes (community, autonomy, and divinity). *Journal of Personality and Social Psychology, 76*, 574-586.
- Schein, C., & Gray, K. (2018). The theory of dyadic morality: Reinventing moral judgment by redefining harm. *Personality and Social Psychology Review, 22*, 32-70.
- Spitznagel, E. L., & Helzer, J. E. (1985). A proposed solution to the base rate problem in the kappa statistic. *Archives of General Psychiatry, 42*, 725-728.
- Trivers, R. L. (1971). The evolution of reciprocal altruism. *The Quarterly Review of Biology, 46*, 35-57.
- Tybur, J. M., Lieberman, D., & Griskevicius, V. (2009). Microbes, mating, and morality: Individual differences in the three functional domains of disgust. *Journal of Personality and Social Psychology, 97*, 103-122.
- Waytz, A., Dungan, J., & Young, L. (2013). The whistleblower's dilemma and the fairness-loyalty tradeoff. *Journal of Experimental Social Psychology, 49*, 1027-1033.