Asymmetric Memory for Harming Versus Being Harmed

Chelsea Helion
Columbia University

Erik G. Helzer
The Johns Hopkins Carey Business School

Suzie Kim
The University of Chicago

David A. Pizarro
Cornell University

Most people have been both the victim and the perpetrator of a moral transgression at some point in their lives; this article asks whether one set of moral experiences is easier to remember than the other, and why. In Study 1, we documented this basic asymmetry, finding that individuals recalled more instances in which they were the victim of a moral transgression than instances in which they were the perpetrator. In Study 2, we found that this asymmetry in memory arises because experiences of being the victim are perceived more negatively than experiences of being the perpetrator. In Studies 3 and 4, we demonstrated the critical role of intent in this asymmetry, finding that victim memories emphasize perpetrator intent to a greater degree than do perpetrator memories (Study 3), and that the memory asymmetry disappeared when individuals recalled unintentional moral violations (Study 4). Finally, in Study 5, we ruled out a potential alternative mechanism for these effects—that of self-protective motivation on the part of perpetrators. We found that the threat associated with the moral violation moderated victim (but not perpetrator) memories, a finding that is inconsistent with a motivational account for perpetrator memories. This research demonstrates that perceived agency shapes emotional experience and autobiographical memory and speaks to the importance of studying morality as it occurs in everyday contexts.

Keywords: morality, autobiographical memory, emotion

Supplemental materials: http://dx.doi.org/10.1037/xge0000691.supp

In 1907, a young boy named Herman Mankiewicz had his bicycle stolen in front of a public library. The memory of that event stuck with him so much that years later, Mankiewicz used his stolen bicycle as inspiration for one of the most famous plot devices in the history of cinema: It was the origin of "Rosebud," the childhood sled of the enigmatic millionaire in *Citizen Kane* (Meryman, 1978). Whether it is stolen bikes, broken hearts, or bruised egos, the harm that a person experiences at the hands of another would appear to have a lasting impact on their memories and the narratives they construct about their lives. Indeed, research has shown that memories of negative life events feature strongly in people's autobiographical narratives (McAdams, 2011), with moral events playing a central role in people's broader identity (Strohminger & Nichols, 2015). What is less understood, however, is the degree to which autobiographical memories for moral transgressions differ depending on one's role in the event: Would the thief from that day in 1907 have remembered stealing the bicycle as vividly as Mankiewicz remembered having it stolen? The present research examined if and why one's role in a moral transgression—as victim or perpetrator—affects one's autobiographical memory for the episode.

At the core of many of the most plaguing problems in interpersonal relationships are asymmetries in the way people understand and explain their own actions and those of others (Pronin, 2007). For example, individuals on both sides of the political aisle blame each other for negotiation failure, spouses attribute asymmetric contributions to housework (Kruger & Gilovich, 1999), and people tend to believe that they know others better than others know them (Pronin, Kruger, Savitsky, & Ross, 2001). A potential source of these different perspectives, and one reason they may be so deeply entrenched, is that individuals' memories for events might differ based upon their role in episode. In the domain of moral transgressions, this would suggest that transgressors and victims might systematically differ in the memories they hold for episodes in which they harmed or were harmed, respectively.

We approach the question of a perpetrator–victim memory asymmetry using the theoretical framework of the moral dyad (Gray, Waytz, & Young, 2012). This framework posits that

many moral transgressions can be described as consisting of an interaction between two individuals: (a) a moral agent (i.e., perpetrator), whose actions and intentions inflict harm on others, and (b) a moral patient (i.e., victim), who experiences pain and suffering as a consequence of an agent's actions. While the victimhood mindset has been explored before (see Baumeister, Stillwell, & Wotman, 1990; Zitek, Jordan, Monin, & Leach, 2010), the moral dyad presents a useful and potentially generative framework for exploring how individuals construct narratives about their own and others' moral transgressions contextually—that is, as a function of their role within a moral event. The theory also affords specific predictions about how individuals may experience and recall moral transgressions as a patient or an agent. Because affect is a critical component of moral patiency (Gray et al., 2012), this suggests that memories for when one has been the victim of another person's actions may be infused with greater emotionality (specifically, negative emotions). On the other hand, because the ability form and act upon plans is critical to moral agency (Gray et al., 2012), this suggests that intentionality will play a critical role in evaluation of and memory for dyadic moral events. Thus, in this research, we explore two possible—and closely related—contributions that dyadic role may play in the predicted asymmetry in moral memory: (a) the perceived negativity of an event and (b) the perceived intentionality of the perpetrator.

Negativity and Intentionality

While ordinary moral judgments have many components, intentionality and negativity are central to these judgments, as they reflect dimensions on which the consequences (negativity) and motivation (intentionality) of moral transgressions are mapped. Sensitivity to negative and harmful moral actions develops early—children as young as 3 months old are sensitive to negative social behavior, such that they avoid "hinderer" characters and approach "helper" characters that they observed in a previous task (Hamlin, Wynn, & Bloom, 2010). Sensitivity to intentionality, as well, emerges early in childhood. By the age of three, children are sensitive to the intentions of an agent when issuing moral condemnation (Josephs, Kushnir, Gräfenhain, & Rakoczy, 2016), and 8-year-old children are able to integrate and act on information about an agent's intentions when evaluating a harmful outcome (Cushman, Sheketoff, Wharton, & Carey, 2013).

On their own, both dimensions also play a prominent role in the study of memory. For example, negative events are more frequently and vividly recalled than neutral events (Payne & Corrigan, 2007), and negativity tends to boost recollective intensity (if not memory accuracy; Phelps & Sharot, 2008). Research has also shown that individuals are better able to recall agents who acted with intentionality than agents who acted accidentally (Fausey & Boroditsky, 2011; Fausey, Long, Inamori, & Boroditsky, 2010), and describing events with language that emphasizes agents' goals and intentions has been shown to improve memory for the episode (Fausey et al., 2010).

In addition to their independent contributions to memory, negativity and intentionality may reinforce one another in moral evaluation. Perceptions of intentionality and negativity are linked, such that intentional harms (e.g., cheating on a partner) are judged as worse and experienced as more negative than

unintentional harms (e.g., forgetting one's wedding anniversary), even when their objective outcomes are the same (Ames & Fiske, 2013; Gray & Wegner, 2008). Moreover, acts that result in worse consequences for others tend to be judged more intentional, even when all other aspects of the behavior are held constant (Alicke, 2000).

Finally, perceptions of both negativity and intentionality are shaped by one's role (as actor or observer) in the event. Actors tend to view their own harmful actions as less intentional than observers do (Feltz, Harris, & Perez, 2012), and observers tend to perceive more intentionality in the same action than do actors themselves (Malle & Knobe, 1997). The moral disengagement literature (Bandura, 1999) suggests that individuals may attempt to distance themselves from their past bad actions by reconstruing their intent or diminishing negative consequences, suggesting that at least some of this process may be motivated by individuals' desire to maintain positive self-views in the face of negative behavior. For all of these reasons, we anticipated that if asymmetries in moral memory exist, they could be traced to differing perceptions of negativity and intentionality held by perpetrators and victims.

Overview of Hypotheses and Studies

Prior research has found that victims experience stronger negative emotions from a moral transgression than do perpetrators (Baumeister et al., 1990). Relative to perpetrators, victims report that their experiences are more emotionally intense and more negative (Baumeister et al., 1990; Stanley, Henne, Iyengar, Sinnott-Armstrong, & Brigard, 2017). Because of the intensity of victim, relative to perpetrator, events we expected the following:

> *Hypothesis 1 (H1):* Individuals will recall more patient memories relative to agent memories over the same time period (memory asymmetry hypothesis).

> *Hypothesis 2 (H2):* Patient autobiographical memories will be more negative than agent autobiographical memories (negativity asymmetry hypothesis).

> *Hypothesis 3 (H3):* Asymmetric negativity will explain the agent/patient memory asymmetry.

After documenting initial evidence in support of these hypotheses in the first two studies, we sought to understand why victim/patient memories are more negative than perpetrator/agent memories. In particular, why do patients remember events more negatively than agents, especially because agents are (by definition) largely responsible for the harm that was inflicted upon patients? Past work suggests that victims are sensitive to the degree of intentionality in perpetrators' offenses and that perceived intentionality enhances the "sting" of a transgression (Gray & Wegner, 2008). This suggests that enhanced negativity in patients' versus agents' memories for moral transgressions may be the result of asymmetric perceptions of agents' intentionality. If patients see agents' behavior as more intentional than agents themselves do, they will likely remember the event as more negative than will agents.

> *Hypothesis 4 (H4):* Patient autobiographical memories will emphasize agent intent more than agent autobiographical

memories (intent asymmetry), explaining the negativity asymmetry.

Finally, in the last study, we tested whether these documented asymmetries are due to perpetrators' self-protective motivations. Acknowledging that one harmed another or acted unethically is threatening to one's moral self-concept (Bandura, 1999; Mazar, Amir, & Ariely, 2008). It is possible that the memory and negativity asymmetries we predicted result from agents downplaying the negativity caused by their actions as a form of self-protection, rather than reflecting an actual negativity difference in agent versus patient experiences.

To examine this explanation, our final study focused on whether the perpetrator–victim memory asymmetry holds equally for acts that are seen as justified versus unjustified. By justified, we mean that the agent acted in a way that they believe was reasonable, but still ended up hurting another individual. Past work has shown that being able to justify one's behavior reduces the self-concept threat associated with acting immorally (Shalvi, Gino, Barkan, & Ayal, 2015). Thus, if the perpetrator–victim memory asymmetry arises out of agents' self-protective motivations to neutralize threat, agents should be freer to report that their memories for justified (i.e., low self-threat) moral transgressions are more negative and more intense than for unjustified (i.e., high self-threat) transgressions. This motivated our final hypothesis:

Hypothesis 5 (H5): Agents memories will be more negative and more fluent when self-protective motivation is low (i.e., for justified behaviors) relative to when it is high (i.e., for unjustified behaviors).

Study 1

The goal of Study 1 was to test our first hypothesis (H1), that autobiographical patient memories would be more accessible than autobiographical agent memories within the same individuals across the same time period.

Method

Participants. Participants were 100 individuals (66 male, $M_{age} = 30.73$, $SD_{age} = 10.4$) from the U.S. who were recruited from Amazon's Mechanical Turk website. For this and all studies, sample size was selected prior to data collection (each study cell having at least 50 observations) based on guidelines for conducting research with unknown effect sizes (Simmons, Nelson, & Simonsohn, 2013). This and all following studies were conducted in accordance with the guidelines established by the institutional review boards of Cornell University and Columbia University, and participants provided informed consent.

Procedure. Following consent, participants were presented with a brief explanation of the moral dyad. Moral agents were defined as "individuals whose intentions and actions bring about harmful events," and moral patients were defined as "individuals who experience feelings and emotions brought about by the moral agent's actions." Following these definitions, participants were given examples of moral agency and patiency:

> Jenny and Elizabeth are waiting in line at a concert. When the doors to the concert hall open, Jenny sees that Elizabeth is in front of her,

and pushes past Elizabeth to get inside first. Elizabeth falls down and cuts her knee.

Participants were then told that the same individual can be a moral agent or a moral patient depending on the situation in which he or she finds him- or herself. They were then presented with another scenario showing an agency/patiency "flip" from the first example:

> Jenny and Elizabeth are close friends. Jenny tells Elizabeth that she and her boyfriend, Mark, have been fighting recently. Elizabeth tells her not to worry, and that everything will be fine. Later that day, Jenny sees Elizabeth and Mark passionately kissing. Jenny turns away and begins crying.

We used the terms *moral agent* and *moral patient* (rather than the more common lay terms of *victim* and *perpetrator*) in this and all studies for two reasons: (a) to be consistent with the terminology used in the theoretical framework (Gray & Wegner, 2008) that served as the basis for this research and (b) to control for any sort of lay theories about what makes someone a victim or perpetrator, as there are reasons to want to avoid being associated with either term.

Dependent variable: Number of moral memories. Following these examples, participants were asked to recall the number of times that they had been a moral patient and a moral agent within the last 6 months on a sliding scale from 0 to ≥10 (i.e., if a participant reported zero patient memories and one agent memory, they would have a score of 0 for patient and 1 for agent). The order in which participants were asked to recall their memories of being an agent or patient was counterbalanced across participants. Participants were also asked to write down a word that they associated with each recalled memory (to ensure that they were recalling concrete memories rather than merely estimating the number of times that they had been a patient or an agent).[1] In this and all studies, participants were given an unlimited amount of time to recall their memories.

Results and Discussion

To test H1, the effect of moral status (i.e., agent/patient status) on the number of recalled memories, we conducted a paired-samples *t* test on the number of agent versus patient memories recalled. As predicted, there was a main effect of moral status on number of recalled memories, such that participants recalled significantly more transgressions in which they were a moral patient ($M = 2.73$, $SD = 2.91$) than a moral agent ($M = 1.72$, $SD = 2.12$) over the 6-month period, $t(99) = 3.16$, $p = .002$, $d = .32$.

These results provide evidence for the asymmetric moral memory hypothesis (H1)—individuals unequally reported moral agent and patient transgressions over the same time period, such that they recalled being victims/patients significantly more frequently than they recalled being perpetrators/agents. However, while this pattern of results is consistent with the existence of a memory bias (in which participants more readily recall one event over the other), it could also be reflective a simpler possibility—that agents may sometimes cause harm to patients without knowing it (whereas the same is less likely for instances of receiving harm

---

[1] In this and in all studies, following the recall task, participants completed a brief demographic survey.

from an agent). For example, a person who cuts in a line absent-mindedly may never know that he has just taken on the role of a moral agent, even though the individuals in the line behind him (the patients) are probably well aware of his inconsiderate behavior. To ensure that this alternative mechanism was not the cause of our documented asymmetry, in the next four studies we focus participants on one specific event recalled from either perspective.

Study 2

In this study, we conceptually replicated H1 and tested a potential explanation for this memory asymmetry: differences in the experienced negativity of the recalled transgression (H2 and H3). Specifically, we predicted that victim memories would be more negative than perpetrator memories (H2), and that this heightened negativity would explain enhanced recall for victim versus perpetrator events (H3).

Method

Participants. Participants were 250 individuals recruited from Amazon's Mechanical Turk website, all residing in the United States. Eleven participants failed to complete the experiment, resulting in a final sample of 239 participants (153 male, $M_{age} = 28.81$, $SD_{age} = 8.81$).

Procedure. Participants were given the same brief explanation of the moral dyad as in Study 1 (along with the same examples). Following these examples, participants were randomly assigned to either the moral agent condition or the moral patient recall condition. In the moral agent condition, participants were presented with the following prompt: "Take a moment and think about a time when you were a *moral agent*, a time when your actions or intentions led to another person being harmed (physically, emotionally, or mentally)." Participants in the moral patient condition were presented with the following prompt: "Take a moment and think about a time when you were a *moral patient*, a time when the actions of another individual led to you being harmed (physically, emotionally, or mentally)."

Dependent variables: Fluency of recall. For this and future studies, we introduced a new dependent variable for recall, fluency. This was necessitated by our decision to limit participants' recall to a single event.

After recalling the event, participants in both conditions were asked how easy it was to recall, on a scale from 1 (*very difficult*) to 7 (*very easy*), and how many sensory details (e.g., what they were seeing, smelling, or feeling) they were able to recall from the event (none, one to two, three to four, five to six, or seven or more). For clarity, these dependent variables were analyzed separately.

Mediator: Event negativity. Following the recall measures, participants rated how negative the event was on a scale from 1 (*not at all negative*) to 7 (*extremely negative*). Note that the negativity question referred to the negativity of the event itself and not the negativity experienced recalling it.

Results and Discussion

Consistent with H1, and replicating the results of Study 1, we found a main effect of moral status on ease of recall, such that participants in the patient condition found it significantly easier to recall a transgression ($M = 5.06$, $SD = 1.61$) than participants in the agent condition ($M = 4.37$, $SD = 1.87$), $t(237) = 3.05$, $p = .003$, $d = .40$. Participants in the patient condition also recalled more sensory details relative to participants in the agent condition, mean difference $= .26$, $t(237) = 2.04$, $p = .042$, $d = .27$.

Confirming H2, participants in the patient condition indicated that the transgression was significantly more negative ($M = 4.89$, $SD = 1.50$) than participants in the agent condition ($M = 4.21$, $SD = 1.56$), $t(237) = 3.48$, $p < .001$, $d = .45$. To test whether negativity drives the relationship between moral status (agent vs. patient) and memory quality for the transgression (H3), we conducted two mediational analyses (Hayes, 2013): one with ease of recall as the dependent variable and one with the number of sensory details recalled as the dependent variable. Consistent with the results reported above, in both analyses, moral status predicted the negativity of the event ($b = -.69$, $SE = .20$, $p < .001$), and negativity of the event predicted both the ease of recall ($b = .50$, $SE = .07$, $p < .001$) and the number of sensory details recalled ($b = .24$, $SE = .04$, $p < .001$). In addition, moral status predicted both the ease of recall of the event ($b = -.69$, $SE = .22$, $p = .003$) and the number of sensory details recalled ($b = -.26$, $SE = .13$, $p = .042$). The correlation between ease of recall and number of sensory details recalled was significant, $r(237) = .44$, $p < .001$, but the magnitude of the correlation suggests that the two variables tapped into distinct aspects of memory quality.

Consistent with evidence for mediation, the relationship between moral status and the ease of recall dropped to nonsignificance ($b = -.36$, $SE = .21$, $p = .088$) when adjusting for differences in negativity of the recalled transgression, and the 5000-iteration bias-corrected bootstrap 95% confidence interval (CI) indicated that the indirect effect through negativity was significant, $a \times b = -.33$, 95% CI [$-.56$, $-.13$] (Figure 1a). Similarly, the relationship between moral status and the number of sensory details recalled dropped to nonsignificance ($b = -.09$, $SE = .12$, $p = .43$) when adjusting for differences in negativity of the recalled transgression, and a 5000-iteration bias-corrected bootstrap 95% CI indicated that the indirect effect through negativity was significant, $a \times b = -.16$, 95% CI [$-.27$, $-.07$] (Figure 1b). Collectively, these results provide evidence for H1, H2, and H3, demonstrating that asymmetric recall for agents and patients was driven by the increased negativity of the experience for patients relative to agents.

Study 3

In Study 3 we sought to better understand the relationship between one's role in a moral transgression, their remembered negativity of the event, and perceptions of perpetrator intent. Specifically, we directly tested the prediction that victim memories emphasize agent intent more than perpetrator memories, which gives rise to asymmetric negativity for victims versus perpetrators (H4).

Method

Participants. Participants were 100 individuals (68 male, $M_{age} = 35.29$, $SD_{age} = 12.21$) recruited from Amazon's Mechanical Turk website (all living in the United States).

a

Negativity

$b = -.69^*$ $b = .47^*$

Moral Status $b = -.36$ (n.s.)   Ease of
(0 = Patient;   ————————————   Recall
1 = Agent)
                          $(b = -.69^*)$

b

Negativity

$b = -.69^*$ $b = .24^*$

Moral Status $b = -.09$ (n.s.)   # of Sensory
(0 = Patient;   ————————————   Details
1 = Agent)
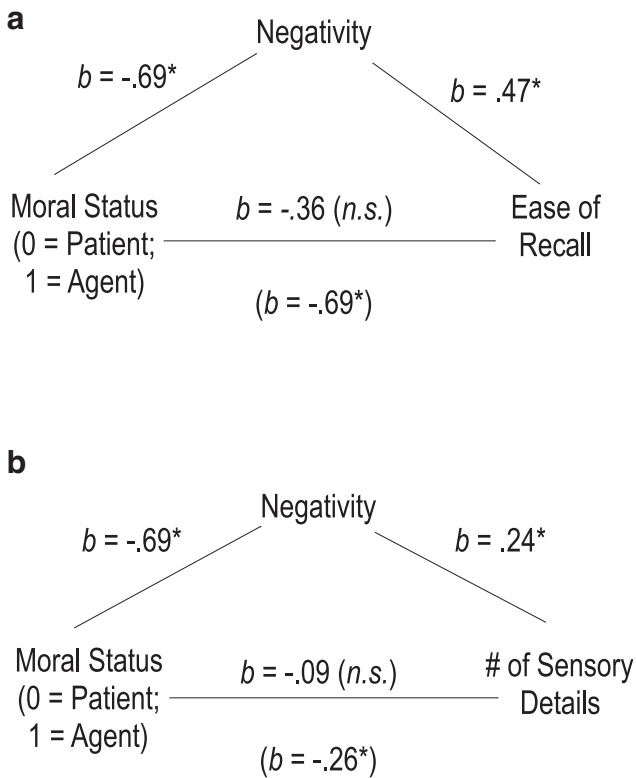 $(b = -.26^*)$

Figure 1. (a) Event negativity mediates the relationship between moral status and ease of recall for a moral transgression. Coefficients represent unstandardized betas. The coefficient in parentheses is the direct effect of moral status on the participants' subjective ease of recall for the moral event. (b) Event negativity mediates the relationship between moral status and the number of sensory details recalled about a moral transgression. All coefficients represent unstandardized betas. The coefficient in parentheses is the direct effect of the independent variable (moral status, i.e., agent or patient) on the number of sensory details recalled. $^*p < .05$.

Procedure. As in Studies 1 and 2, participants were again given a brief explanation, with examples, of the moral dyad. Participants were then randomly assigned to either the moral agent condition or the moral patient condition and viewed the same prompts as those used in Study 2.

Dependent variable: Event negativity. Following memory recall, participants rated how negative the event was on a scale from 1 (*not at all negative*) to 7 (*extremely negative*).

Mediating variable: Agent intent. Participants then answered the following questions to measure agent intent:

> When you were a moral agent/moral patient, how much did you intend to harm another person? How much do you think that the moral agent intended to harm you (physically, emotionally, or mentally)? That is, how intentional was your behavior/how intentional do you think the actions of the moral agent were? (Ratings made on a scale from 1 = *not at all intentional* to 7 = *extremely intentional*)

> When you were a moral agent, how in control were you of your behavior? How in control do you think the moral agent was of his or her behavior? (Ratings made on a scale from 1 = *I/they was/were not at all in control of my/their behavior* to 7 = *I/they was/were extremely in control of my/their behavior*)

> When you were a moral agent, how planned was your behavior? How planned do you think the moral agent's behavior was? (Ratings made on a scale from 1 = *not at all planned* to 7 = *extremely planned*).

> When you were a moral agent, how responsible were you for the event that occurred? How responsible was the moral agent for the event that occurred? (Ratings made on a scale from 1 = *not at all responsible* to 7 = *extremely responsible*)

Past research (Malle & Knobe, 1997) indicated that folk concepts of intentionality tend to reflect aspects of each of these dimensions. The four items were averaged into a composite of agent intent (Cronbach's $\alpha = .75$).

Results and Discussion

Consistent with H2, and replicating the results of Study 2, we found a main effect of moral status on negativity, such that transgressions in which participants were in the role of moral patient ($M = 5.47$, $SD = 1.42$) were significantly more negative than transgressions in which participants were in the role of moral agent ($M = 4.71$, $SD = 1.40$), $t(98) = 2.71$, $p = .008$, $d = .54$. Consistent with H4, we also found the predicted discrepancy in perceived intent, such that individuals in the moral patient condition ($M = 5.35$, $SD = 1.31$) perceived significantly more intentionality than individuals in the agent condition ($M = 3.82$, $SD = 1.10$), $t(98) = 6.31$, $p < .001$, $d = 1.26$.

To directly test whether the relationship between moral status and event negativity was due to differences in perceived agent intent (H4), we conducted a mediational analysis (Hayes, 2013). Consistent with the results reported above, we found that moral status predicted both event negativity ($b = -.76$, $SE = .28$, $p = .008$) and agent intent ($b = -1.53$, $SE = .24$, $p < .001$). In addition, there was a significant relationship between perceived intent and event negativity ($b = .39$, $SE = .09$, $p < .001$), and the relationship between moral status and event negativity dropped to nonsignificance ($b = -.23$, $SE = .32$, $p > .40$) when adjusting for differences in agent intent. A 5,000-iteration, bias-corrected bootstrap confirmed H4, indicating that the indirect effect through agent intent was significant, $a \times b = -.53$, 95% CI [−.94, −.19] (see Figure 2).

These results lend robust support to H2 and H4. Individuals found transgressions more negative when they were in the role of

Agent Intent

$b = -1.53^*$ $b = .39^*$

Moral Status
(0 = Patient; $b = -.23$, (n.s.)   Negativity
1 = Agent)   ————————————
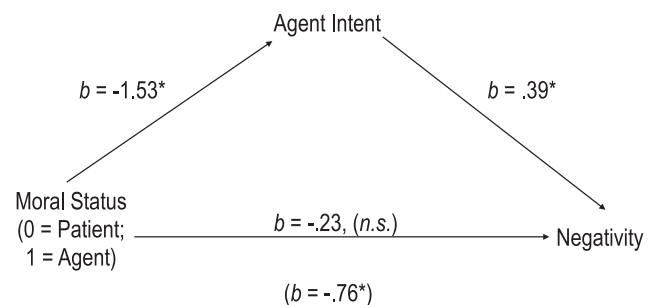 $(b = -.76^*)$

Figure 2. Perceived agent intent mediates the relationship between moral status (i.e., agent or patient) and participants' subjective reports of event negativity. All coefficients are unstandardized betas. The coefficient in parentheses is the direct effect of moral status on participants' subjective reports of negativity. $^*p < .05$.

moral victim compared to when they were in the role of moral perpetrator, and much of this effect was due to asymmetric perceptions of intentionality across the two conditions. Victims perceived the perpetrators who harmed them to be acting with greater intentionality than perpetrators perceived in their own behavior—and this intensified the negativity of the experience for victims.

One concern that could be raised about the present study is that the instructions given prior to recall may have biased the results toward H4. In our instructions to participants, agency was characterized by actions and intentions, whereas patiency was characterized by emotional experience. Note, however, that these definitions were given to all participants and that reports of intent were always focused on the moral agent, regardless of whether participants were recalling the memory from the role of the agent or patient. Thus, although our definition of agency stipulates intentionality, this definition would apply equally to participants who recall the memory from the agent or patient perspective. In other words, the stipulation of intentionality would not yield an agent/patient asymmetry in ratings of the agent's intentionality. However to address this concern more fully, in Study 4, we experimentally manipulated agent intent to observe its impact on both recalled negativity and memory intensity.

Study 4

The aim of Study 4 was twofold: (a) to replicate H4 in a fully experimental context by manipulating agent intent and (b) to simultaneously account for moral status, negativity, and intent within the same model as predictors of the moral memory asymmetry. By constraining agents and patients to recall instances of either intentional or unintentional transgressions, we can examine whether or not the memory asymmetry holds when individuals in each role recall events matched on perceived intentionality. If there is no difference between agents and patients in negativity and memory recall when intentionality is low, but the asymmetry remains when intentionality is high, this would be evidence that the perceived presence (or absence) of agent intent differentially impacts agents and patients with regard to negativity and memory recall.

Method

Participants. Participants were 200 individuals residing in the United States (143 male, $M_{age} = 28.38$, $SD_{age} = 8.22$) recruited from Amazon's Mechanical Turk website.

Procedure. As in Studies 2 and 3, participants were given an explanation of the moral dyad and the same examples. Participants were then randomly assigned to one condition in a 2 (Moral Status: agent or patient) $\times$ 2 (Agent Intent: intentional or accidental) between-subjects design. Moral status was manipulated as in Studies 2 and 3. Agent intent was manipulated by having participants recall a time when the harm caused by the agent (or themselves, if they were in the agent condition) was either accidental or intentional.[2] Participants in the accidental condition saw one of the following prompts:

> *Patient condition:* Take a moment and think about a time when you were a *moral patient* because a moral agent acted accidentally. That is, a time when the accidental actions of another individual led to you being harmed (physically, emotionally, or mentally).

> *Agent condition:* Take a moment and think about a time when you were accidentally a *moral agent*, that is, a time when your actions or intentions accidentally led to another person being harmed (physically, emotionally, or mentally).

And participants in the intentional condition saw one of the following prompts:

> *Patient condition:* Take a moment and think about a time when you were a *moral patient*, because a moral agent acted intentionally. That is, a time when the intentional actions of another individual led to you being harmed (physically, emotionally, or mentally).

> *Agent condition:* Take a moment and think about a time when you were intentionally a *moral agent*, a time when you knew that your actions or intentions would lead to another person being harmed (physically, emotionally, or mentally).

Dependent variable: Memory quality. Following memory recall, participants in all four conditions were asked how negative the event was, followed by three questions that assessed memory fluency: how easy the memory had been to recall, how vivid their memory was for the event, and how confident they were in their memory for the event. These new measures provided a more robust measure of recollective experience, and were averaged into an index of memory quality (Cronbach's $\alpha = .86$). All questions were evaluated on a seven-point scale from 1 (*not at all*) to 7 (*extremely*). Following these measures, participants completed a brief demographic survey.

Results and Discussion

One participant failed to complete the experiment, leaving us with 199 participants.

Intentionality and negativity. We subjected ratings of negativity to a 2 $\times$ 2 between-subjects analysis of variance (ANOVA). The ANOVA returned a significant main effect of moral status, $F(1, 195) = 5.61$, $p = .019$, $\eta_p^2 = .028$, no main effect of intentionality, $F(1, 195) = 2.67$, $p > .10$, and the predicted status $\times$ intentionality interaction, $F(1, 195) = 15.10$, $p < .001$, $\eta_p^2 = .072$. Consistent with H4 and the results of Study 3, a post hoc Tukey's test exploring this interaction revealed that patients experienced more negativity ($M = 5.33$, $SD = 1.34$) than did agents ($M = 4.12$, $SD = 1.28$), $p < .001$, for intentional harms. For accidental harms, where intentionality was absent, there was no significant difference between patients ($M = 4.26$, $SD = 1.45$) and agents ($M = 4.57$, $SD = 1.40$) $p > .60$, in experienced negativity (see Figure 3).

Testing the moderated mediation model. In the next analysis, we model the relationship between moral status, intentionality, negativity, and moral memory, collectively accounting for H1–H4. Based on the preliminary results above, we ran a moderated

---

[2] One question may be whether or not unintentional transgressions constitute moral violations. Prior research has found that individuals (and even children) hold actors responsible for negative side effects of their actions, even if the side effects were brought about unintentionally (Leslie, Knobe, & Cohen, 2006), and individuals use outcome information (rather than relying solely on intent) to assess blameworthiness (Cushman, 2008). Thus, regardless of whether unintended transgressions constitute actual (i.e., normative) moral violations, in many real world cases, perceivers render moral judgments on them as if they were immoral.
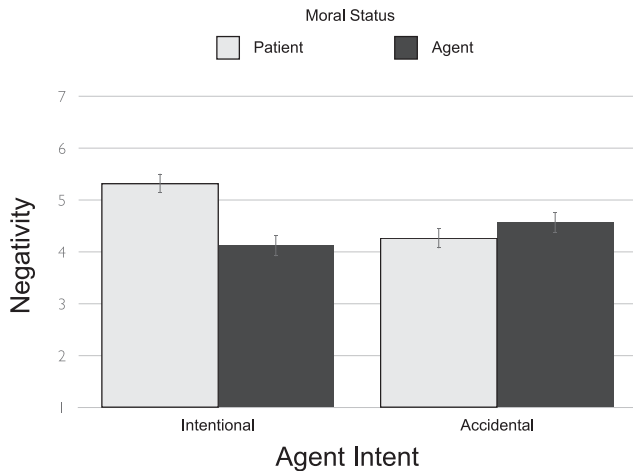
Figure 3. The interaction between moral status (agent, patient) and agent intent (intentional, accidental) predicting reported event negativity. Patients report that the event is more negative than agents when the agent acted intentionally, but not when the agent acted accidentally (error bars indicate ±1 *SE* from the condition mean).

mediation model, testing whether asymmetric moral memory for agents and patients (H1) can be traced to asymmetric negativity for those events (H2 and H3), but only when acts are judged as intentional (H4). We did not anticipate negativity or memory asymmetries for accidental transgressions, because these acts, by definition, lack clear intentionality.

The mediation package in R (Tingley, Yamamoto, Hirose, Keele, & Imai, 2014) was used to estimate the model coefficients and conditional direct and indirect effects. As shown in Figure 4, the effect of moral status on negativity was contingent upon agent intent, as seen by the significant interaction between moral status and intent in the model of negativity ($b = -1.51$, $SE = .39$, $p < .001$). There was also a significant indirect effect of moral status on memory quality when the agent acted intentionally ($a \times b = -.58$, 95% CI [$-.98$, $-.25$]) but not when the agent acted accidentally ($a \times b = .11$, 95% CI [$-.09$, $.37$]). Thus, consistent with our predictions, negativity mediated the effect of moral status on memory intensity when the harm was intentional, but not when the harm was accidental (see Figure 4). Taken together, these results indicate that perceived agent intentionality is an important causal contributor to the memory asymmetry in that it determines whether the negativity associated with a moral transgression will translate to increased memory intensity.

Study 5

In Study 5 we sought to clarify the role of self-protective motivation in establishing the negativity and memory asymmetries between perpetrators and victims. From a motivated cognition perspective, the asymmetries we have documented could be aimed at threat reduction. To protect the self against threat, individuals may selectively forget their own versus others' (intentional) moral transgressions, and/or downplay their severity through diminished intent or negativity. This logic motivated H5, which we test here.

To test the impact of self-protective motives on the moral memory asymmetry, we created a high-threat and low-threat con-

dition based on recent work on moral justification. The ability to generate a justification for one's behavior has been shown to reduce self-concept threats associated with acting immorally (Shalvi et al., 2015), and to impact assessments of perpetrator guilt and victim distress (McGraw, 1987). Thus, if motivated processes are at play, we would expect agents' memories for moral transgressions to be affected by the degree to which they have justified their actions to themselves. Specifically, H5 holds that the agent/patient asymmetries for negativity and memory quality will hold more strongly for unjustified transgressions, which pose greater threat to agents, than for justified transgressions, which pose less threat.

Method

Participants were 160 individuals residing in the United States (68 male, $M_{age} = 40.13$, $SD_{age} = 11.34$) recruited from Amazon's Mechanical Turk website. Sample size was decided by selecting a stopping date for collection a priori. Two participants were excluded for writing responses that were unintelligible; four participants were excluded for not following instructions to recall an event within the past year, leaving us with 154 participants in the final sample.

Procedure. As in Studies 2–4, participants were given an explanation of the moral dyad and the same examples. Participants were then randomly assigned to one condition in a 2 (Moral Status: agent or patient) × 2 (Act Justifiability: justified or unjustified)
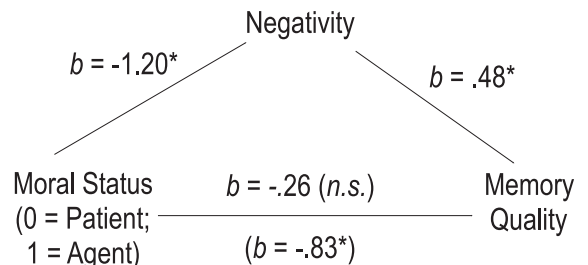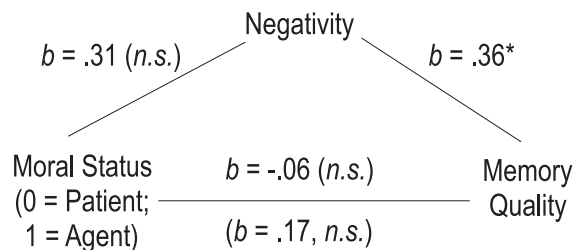

Figure 4. When agents acted intentionally, event negativity mediated the relationship between moral status and memory quality for the moral event. This was not the case when the agent acted accidentally. All betas are unstandardized coefficients. The coefficient in parentheses is the direct effect of moral status on participants' subjective reports of negativity. * $p < .05$.

between-subjects design. Moral status was manipulated as in Studies 2–4. Act justifiability was manipulated by having participants recall a time when the harm caused by the agent (or themselves, if they were in the agent condition) was either justifiable or unjustifiable. Participants in all conditions saw the following prompt with a definition of justified and unjustified within the context of moral memory:

> Most of us have experienced being both moral agents and moral patients, sometimes involving actions that seem more or less justified than others. By *justified*, we mean performing an action for a sufficient or legitimate reason. By *unjustified*, we mean performing an action for an insufficient or illegitimate reason. Sometimes we hurt others (and vice versa) in ways that are *justified* in order to get our needs met, and sometimes we may lash out in the heat of the moment in ways that are *unjustified*.

Participants in the unjustified condition then recalled either a time that there were a patient due to another person's unjustifiable actions (patient condition), or when they were an agent to another person via unjustifiable actions (agent condition). Participants in the justified condition recalled a justifiable action on the part of the agent (patient condition), or themselves (agent condition)[3]. Participants in all conditions were instructed to write a short phrase that they associated with the recalled event, and to limit their search to events from the past year. We expanded the time frame to a year (rather than the 6 months used in Study 1), as we were constraining the type of memory participants were recalling (patient or agent and justified or unjustified), and felt that giving participants a larger time window increased the likelihood of being able to recall a memory that fit the study criteria.

Dependent variable: Event negativity. Participants rated how negative the event was on a scale from $1 = $ *not at all negative* to $7 = $ *extremely negative*.

Dependent variable: Fluency of recall. As in Studies 2 and 4, participants again rated how easy the memory had been to recall (on a scale from $1 = $ *very difficult* to $7 = $ *very easy*), and how many sensory details (none, one to two, three to four, five to six, seven or more) they could recall from the event.

Manipulation check: Act justifiability. Following these measures, participants were asked to indicate their agreement with three statements meant to assess the effectiveness of the justifiability manipulation: "Most people would act like the moral agent/I did in this situation," "Most people would agree that the moral agent's/my actions were justified in this situation," "The moral agent/I acted the way that they/I did for a sufficient or legitimate reason." All agreement assessments were made on a 7-point scale from $1 = $ *strongly agree* to $7 = $ *strongly disagree*. The justifiability questions had high reliability (Cronbach's $\alpha = $ .94) and were averaged to create an index of justifiability. In addition to these measures, we also collected several exploratory measures.[4]

Results

Justifiability manipulation check. Our manipulation of act justifiability was successful: Ratings of justifiability (reverse-coded) were higher among participants in the justified condition, $M = 5.05$, $SD = 1.58$, than among participants in the unjustified condition ($M = 2.82$, $SD = 1.61$), $t(152) = 8.63$, $p < .001$, $d = 1.4$.[5]

Event negativity. To test H5, we ran a two-way ANOVA to examine the relationship between justifiability (between subjects: justified, unjustified) and moral status (between subjects: patient, agent) on event negativity. We found a marginal main effect of moral status on negativity, $F(1, 150) = 2.84$, $p = .094$, $\eta_p^2 = .019$. We also found significant main effect of act justifiability on negativity, $F(1, 150) = 27.30$, $p < .001$, $\eta_p^2 = .15$, such that justified harms were experienced as less negative than unjustified harms. These main effects were qualified by a significant Moral Status $\times$ Act Justifiability interaction, $F(1, 150) = 3.88$, $p = .05$, $\eta_p^2 = .025$ (see Figure 5).

A post hoc Tukey's test exploring this interaction found that the memory asymmetry was in part a function of the justifiability of the moral transgression. However, contrary to H5, the justifiability of the act was more impactful for patients than for agents. For unjustified transgressions, patients ($M = 5.76$, $SD = 1.13$) experienced greater negativity than agents ($M = 4.92$, $SD = 1.44$) $p = .05$. However, for justified transgressions there was no significant difference between patients ($M = 4.34$, $SD = 1.39$) and agents ($M = 4.36$, $SD = 1.27$) for event negativity, $p = .99$. Furthermore, there was no difference in experienced negativity across the two agent conditions, mean difference $= .56$, $p = .361$. Instead, there was a significant difference in experienced negativity across the two patient conditions, mean difference $= -1.41$, $p < .001$. That is, participants in the patient, not the agent, condition drove the interaction (see Figure 5).

Fluency of recall. The correlation between number of sensory details recalled and ease of recall was significant, but modest, $r(152) = .28$, $p < .001$. We therefore analyze these dependent measures separately.

We found no main effect of moral status on the number of details recalled, $F(1, 150) = 1.73$, $p = .19$. We also found no main effect of act justifiability on the number of details recalled, $F(1,$

---

[3] There was a typo in one section of the instructions for the agent condition, to recall "a times" rather than "a time" in which they had acted as an agent justifiably or unjustifiably. However in other instructions for the condition, it is clear that agents should recall one memory rather than multiple memories (including on the page with the memory recall prompt) and we do not believe that this typographical error impacted our results.

[4] Participants indicated the nature of their primary relationship with the moral agent or patient (e.g. current romantic relationship, former romantic relationship, family member), reported how close of a relationship they had with the moral agent or patient before the incident occurred, and how close of a relationship they currently had with the moral agent or patient. Participants also completed the Guilt and Shame Proneness Scale (Cohen, Wolf, Panter, & Insko, 2011) and a brief demographic survey. Given that we did not have specific hypothesis regarding the nature of guilt proneness or relationship closeness on the dependent variables of interest, these exploratory analyses are not the primary focus of the manuscript and are included in online supplementary materials.

[5] Planned comparisons indicated that there were no differences between agents ($n = 26$, $M = 3.21$, $SD = 1.56$) and patients ($n = 45$, M = 2.59, $SD = 1.62$) in the unjustifiable condition in terms of agreement that the agent acted unjustifiably, $t(69) = 1.56$, $p = .12$. There was, however, a difference between agents ($n = 33$, $M = 5.49$, $SD = 1.46$) and patients ($n = 50$, $M = 4.75$, $SD = 1.61$) in the justifiable condition, such that agents saw their actions as more justified than did patients, $t(81) = 2.13$, $p = .036$, $d = .47$.
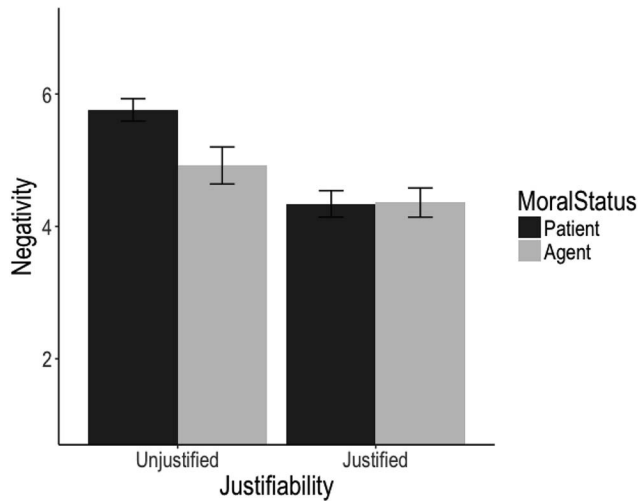
Figure 5. Interaction between moral status and justifiability of the action predicting event negativity. Although there was an asymmetry in experienced negativity between patients and agents when the action is viewed as unjustified, there was no difference in negativity when the action was considered justified (error bars indicate ±1 *SE* from the condition mean).

150) = 1.03, p = .31. We did, however, find the predicted two-way interaction between moral status and justifiability on the number of details recalled, $F(1, 150) = 6.94$, $p = .009$, $\eta_p^2 = .044$, which mirrored results obtained for event negativity. Post hoc Tukey's tests examining the nature of this interaction found that in the unjustified condition patients recalled more details than agents, mean difference = .65, p = .027. In the justified condition, however, patients and agents recalled the same degree of details, mean difference = .16, p = .85. We found no difference in recall fluency across the agent conditions, mean difference = .36, p = .45. Paralleling the negativity results, we found a marginal difference across the two patient conditions, mean difference = .46, p = .083.

In addition, we found a main effect of justifiability on ease of recall, $F(1, 150) = 7.60$, $p = .007$, $\eta_p^2 = .047$, indicating that it was easier for participants to recall unjustified, relative to justified, memories. We found no main effect of moral status, $F < 2$, and there was no significant two-way interaction between moral status and justifiability on ease of recall, $F < 2$.

Taken together, these results clarify two aspects of the preceding studies. First, our results failed to provide support for H5, and cast doubt on the possibility that the asymmetries observed in Studies 1–4 are the result of individuals deploying motivated, self-protective processes to ward off threat. If they were, the justifiability manipulation would have affected agents' recollection; instead, we observed just the opposite. The fact that justifiability affected patients' recollection suggests instead that patients' enhanced memory for moral transgressions may be part of a broader sense-making process, in which patients seek to understand why they were victims at the hands of another person. Unjustified transgressions may loom larger in memory for patients precisely because they have not fully made sense of the episode; justified transgressions may fade from patients' memories (as they do for

agents) because the process of justifying another person's behavior indicates that one has adequately made sense of it.

General Discussion

Looking back on their lives, most people can recall episodes in which their actions harmed others and episodes in which others' actions harmed them. Such memories are part of the raw material from which people build broader life narratives (McAdams, 2011), and come to understand themselves as moral beings (Strohminger & Nichols, 2015). The aim of the present research was to investigate people's memories for such events—in particular, to examine whether people showed memory asymmetries for the times they acted as perpetrators versus victims of moral transgressions, and if so, why. Four out of five hypotheses tested in five studies received strong support, suggesting that victim memories are more frequently and more easily recalled than perpetrator memories, and that this is due to asymmetric negativity experienced by perpetrators and victims at the time of the transgressions.

In Studies 1 and 2, we found that memories of moral victimhood were more easily and fluently recalled than memories of being a moral perpetrator (H1), and that this asymmetry was explained by differences in the experienced negativity of the event, such that victim memories were more negative than perpetrator memories (H2 and H3). Studies 3 and 4 focused on the role that the intentionality of an agent plays in establishing this asymmetry in recall. Study 3 found that differences between agents and patients in perceptions of agents' intentions mediated the established relationship between moral status and negativity—victims experienced moral transgressions more negatively because the actors involved in one's own victimhood were perceived as having acted with greater intentionality (H4). Study 4 established a causal role for intentionality, revealing that asymmetric negativity and moral memory held for intentional acts only. Finally, in Study 5, we examined the role that self-protective motivation may play in establishing the moral memory asymmetry by manipulating the moral justifiability of the moral transgression. Contrary to H5, we found that the asymmetric memory effect disappeared when agents and patients recalled events in which the agent acted in a way that was perceived as justified, but that this was due to differences in how patients (but not agents) were impacted by justifiability concerns.

Given this pattern of results, we postulate that the moral memory asymmetry documented across these studies is the result of two (potentially additive) explanations: (a) differences in perceived negativity and (b) differences in the extent to which victims and perpetrators are able to make sense of the event. We found that the experience of being a victim was often more negative than that of being a perpetrator, and that victims, relative to perpetrators, were much more sensitive to differences in agent intent and act justifiability. Moreover, victims' enhanced memory for moral transgressions was attenuated when victims saw the perpetrator's behavior as justified, suggesting that the memory asymmetry aids in sense-making efforts on the part of victims. From the perspective of victims, unjustified moral transgressions are often those for which there is inadequate explanation—and it was for these types of memories that the negativity and moral memory asymmetries were observed. The tendency for events that feel unfinished to remain active in memory has been well documented (Zeigarnik, 1938;

Bargh, Gollwitzer, Lee-Chai, Barndollar, & Trötschel, 2001; Savitsky, Medvec, & Gilovich, 1997), and the memory asymmetry observed in our studies may reflect the extent to which these events feel incomplete and in need of further processing from the perspective of the victim relative to the perpetrator. It is also possible that as victims attempt to make sense of an event, they engage in more rehearsal of the memory, resulting in improved recall. This explanation is consistent with the literature on rehearsal as a process that enhances memory consolidation and retrieval (Waugh & Norman, 1965; Davachi, Maril, & Wagner, 2001), and is a promising avenue for future research.

Limitations and Future Directions

Although our data lend robust support to the existence of a moral memory asymmetry and provide a detailed picture of why and for whom the asymmetry arises, a number of questions are open for future research. First, the degree to which the moral memory asymmetry reflects differences in memory construction versus memory accessibility remains unclear. Put a different way, do our results suggest that people actually experience victim events with greater frequency, greater negativity, and greater agent intent than perpetrator events as a matter of lived experience (which is then reflected in autobiographical memory; the memory-construction hypothesis), or are patient memories featuring heightened negativity and agent intent simply more accessible in memory than agent memories, regardless of the actual frequency and experience of such events in one's life (the accessibility hypothesis)? Our present studies are not sensitive enough to disentangle these precise processes; however, a deeper understanding of this issue would shed light on the degree to which constructive versus reconstructive processes are involved in people's understanding of themselves as moral beings. We should note that these explanations are not at odds with one another—it is possible, and in our view quite likely, that the memory asymmetry we have observed here is the product of both actual differences between perpetrators and victim experiences in vivo as well as downstream processes that result in heightened accessibility for victim relative to perpetrator memories.

A second limitation is that in both our studies and in the extant literature, the roles of agent and patient are treated as fixed—an individual in a moral event is designated clearly as one or the other. However, in the context of everyday moral transgressions the roles of agent and patient are often more ambiguous and fluid. Consider the context of a long-term relationship. One individual may harm another because they feel that the other harmed them first, making the clean division into victim and perpetrator as modeled in this work perhaps less clear. Moreover, past work has shown that where one draws the line between agents and patients in autobiographical memory can be affected by broader contextual and motivational factors. For example, Wohl and Branscombe (2008) found that Jewish Canadians who were reminded of the Holocaust (vs. Jewish Canadians in a control condition) felt that Israel's actions against Palestinians were more merited and that Israel bore less responsibility. In this case, reminders of past victimhood affected individuals' perceptions of themselves (and their group) as moral agents in the present context. With respect to both considerations, our work points out that it is the subjective assessment of oneself as an agent or patient in a particular context

that matters for moral memory. Thus, if both relationship partners (or both parties on either side of a geopolitical conflict) see themselves as victims of the other's actions, we would expect on the basis of our data that both would show enhanced memory and negativity for the event, regardless of the "true" cleaving of agent and patient for that particular episode.

Turning to other literature, the present results may seem somewhat surprising in light of relevant research on guilt and shame. Prior work has found that individuals can experience guilt and discomfort following moral transgressions (Baumeister, Reis, & Delespaul, 1995; McGraw, 1987) and that shame-eliciting events occupy a central role in one's autobiographical memory (Pinto-Gouveia & Matos, 2011). Individuals report high levels of guilt when recalling events in which they betrayed or were rude to a relationship partner, family member, or close friend, but that in general, individuals tend to experience relatively low levels of guilt in their daily life (Baumeister et al., 1995). Perpetrators also tend to report higher levels of guilt when the event was accidental relative to when it was intentional (McGraw, 1987). While we did not find a difference in reported negativity when agents were recalling accidental versus intentional events (Study 4), this may be because we did not ask about guilt specifically, but rather negativity more generally. Future work should examine the role that perpetrator guilt may play in establishing or mitigating the memory asymmetry between moral patients and agents.

Given that participants in the agent condition were less likely to recall events in which they had acted intentionally, it's possible these may have also been events that were less characterized by guilt or shame; however, the results of Study 4, in which there were no appreciable differences in memory for agents who recalled intentional versus unintentional moral transgressions, seem to speak against this possibility. It seems more likely that our results diverge from what might be expected from the guilt and shame literature because we had participants recall moral transgressions that were relatively mild in intensity. Turning to the results of Study 1 (in which participants were instructed to recall both agent and patient memories using a one-word cue), many of the most common words listed by agents were common moral transgressions—for example, *fight*, *cheated*, *gossip*, *breakup*, and *lying*. When in the patient condition, they recalled similar events: *ignored*, *lied*, *fight*, *yelling*, and *cheating*. The choice to focus on more commonplace moral transgressions was a deliberate one, as the primary goal of our research was to understand the psychological processes that comprise and create the everyday moral landscape. That said, it's possible that the asymmetries observed across agents and patients in our study may not hold for particularly severe or impactful moral transgressions, and even may reverse at the extremes. The case of moral injury—the emotional and psychological trauma that results from having acted in a way that violates one's moral rules and damages one's moral self-concept (Maguen & Litz, 2012)—may be one context in which the asymmetry observed in our studies is attenuated or even reversed. While addressing these questions is outside the bounds of the present research, we believe that it is an important avenue for future study.

Finally, it remains unknown whether a memory asymmetry would hold for positive moral events—one in which the moral agent is the benefactor of a moral event, and the moral patient is a beneficiary. Individuals tend to have more recent memories of

times when they helped as compared to harmed others (Escobedo & Adolphs, 2010), and an accessibility bias for positive memories of the self over negative ones is consistent with a motivational account of autobiographical moral memory. However, it is unclear if the memory asymmetry observed here would hold for benefactor/beneficiary events—receiving help is certainly a more positive experience than not receiving it, but needing help in the first place is perhaps not.

In total, this research has illuminated robust biases in how moral events are perceived and recalled, resulting in a memory asymmetry biased toward the recall of victim relative to perpetrator memories. This suggests that not only is the experience of each role different as it occurs, but that it continues to be different as individuals look back and construct the narrative of their lives. Thus, while people may often talk, text, or even write Academy-Award winning screenplays about the bad things that have happened in their lives, they may be notably silent on the bad things that they have caused in the lives of others.

Context

One of the primary goals of this research was to examine how individuals construct narratives of the moral events in their lives, and how the accessibility and content of these narratives differs as a function of the role one plays in a moral event. This work was meaningfully influenced by theories about dyadic morality (see Gray et al., 2012; Gray & Wegner, 2008), and builds on previous work from the authors that examined biases in memory based on agent intent (Pizarro, Laney, Morris, & Loftus, 2006), the role of emotion in moral decision-making (Helion & Pizarro, 2015; Helion & Ochsner, 2018), and perceptions of agency in self and other (Helzer & Dunning, 2012).

References

Alicke, M. D. (2000). Culpable control and the psychology of blame. *Psychological Bulletin, 126,* 556–574. http://dx.doi.org/10.1037/0033-2909.126.4.556

Ames, D. L., & Fiske, S. T. (2013). Intentional harms are worse, even when they're not. *Psychological Science, 24,* 1755–1762. http://dx.doi.org/10.1177/0956797613480507

Bandura, A. (1999). Moral disengagement in the perpetration of inhumanities. *Personality and Social Psychology Review, 3,* 193–209. http://dx.doi.org/10.1207/s15327957pspr0303_3

Bargh, J. A., Gollwitzer, P. M., Lee-Chai, A., Barndollar, K., & Trötschel, R. (2001). The automated will: Nonconscious activation and pursuit of behavioral goals. *Journal of Personality and Social Psychology, 81,* 1014–1027. http://dx.doi.org/10.1037/0022-3514.81.6.1014

Baumeister, R. F., Reis, H. T., & Delespaul, P. A. E. G. (1995). Subjective and experiential correlates of guilt in daily life. *Personality and Social Psychology Bulletin, 21,* 1256–1268. http://dx.doi.org/10.1177/01461672952112002

Baumeister, R. F., Stillwell, A., & Wotman, S. R. (1990). Victim and perpetrator accounts of interpersonal conflict: Autobiographical narratives about anger. *Journal of Personality and Social Psychology, 59,* 994–1005. http://dx.doi.org/10.1037/0022-3514.59.5.994

Cohen, T. R., Wolf, S. T., Panter, A. T., & Insko, C. A. (2011). Introducing the GASP scale: A new measure of guilt and shame proneness. *Journal of Personality and Social Psychology, 100,* 947–966. http://dx.doi.org/10.1037/a0022641

Cushman, F. (2008). Crime and punishment: Distinguishing the roles of causal and intentional analyses in moral judgment. *Cognition, 108,* 353–380. http://dx.doi.org/10.1016/j.cognition.2008.03.006

Cushman, F., Sheketoff, R., Wharton, S., & Carey, S. (2013). The development of intent-based moral judgment. *Cognition, 127,* 6–21. http://dx.doi.org/10.1016/j.cognition.2012.11.008

Davachi, L., Maril, A., & Wagner, A. D. (2001). When keeping in mind supports later bringing to mind: Neural markers of phonological rehearsal predict subsequent remembering. *Journal of Cognitive Neuroscience, 13,* 1059–1070. http://dx.doi.org/10.1162/089892901753294356

Escobedo, J. R., & Adolphs, R. (2010). Becoming a better person: Temporal remoteness biases autobiographical memories for moral events. *Emotion, 10,* 511–518. http://dx.doi.org/10.1037/a0018723

Fausey, C. M., & Boroditsky, L. (2011). Who dunnit? Cross-linguistic differences in eye-witness memory. *Psychonomic Bulletin & Review, 18,* 150–157. http://dx.doi.org/10.3758/s13423-010-0021-5

Fausey, C. M., Long, B. L., Inamori, A., & Boroditsky, L. (2010). Constructing agency: The role of language. *Frontiers in Psychology, 1,* 162. http://dx.doi.org/10.3389/fpsyg.2010.00162

Feltz, A., Harris, M., & Perez, A. (2012). Perspective in intentional action attribution. *Philosophical Psychology, 25,* 673–687. http://dx.doi.org/10.1080/09515089.2011.627540

Gray, K., Waytz, A., & Young, L. (2012). The moral dyad: A fundamental template unifying moral judgment. *Psychological Inquiry, 23,* 206–215. http://dx.doi.org/10.1080/1047840X.2012.686247

Gray, K., & Wegner, D. M. (2008). The sting of intentional pain. *Psychological Science, 19,* 1260–1262. http://dx.doi.org/10.1111/j.1467-9280.2008.02208.x

Hamlin, J. K., Wynn, K., & Bloom, P. (2010). Three-month-olds show a negativity bias in their social evaluations. *Developmental Science, 13,* 923–929. http://dx.doi.org/10.1111/j.1467-7687.2010.00951.x

Hayes, A. F. (2013). *Methodology in the social sciences. Introduction to mediation, moderation, and conditional process analysis: A regression-based approach.* New York, NY: Guilford Press.

Helion, C., & Ochsner, K. N. (2018). The role of emotion regulation in moral judgment. *Neuroethics, 11,* 297–308. http://dx.doi.org/10.1007/s12152-016-9261-z

Helion, C., & Pizarro, D. A. (2015). Beyond dual-processes: The interplay of reason and emotion in moral judgment. In J. Clausen & N. Levy (Eds), *Handbook of neuroethics* (pp. 109–125). Dordrecht, the Netherlands: Springer.

Helzer, E. G., & Dunning, D. (2012). Why and when peer prediction is superior to self-prediction: The weight given to future aspiration versus past achievement. *Journal of Personality and Social Psychology, 103,* 38–53. http://dx.doi.org/10.1037/a0028124

Josephs, M., Kushnir, T., Gräfenhain, M., & Rakoczy, H. (2016). Children protest moral and conventional violations more when they believe actions are freely chosen. *Journal of Experimental Child Psychology, 141,* 247–255. http://dx.doi.org/10.1016/j.jecp.2015.08.002

Kruger, J., & Gilovich, T. (1999). "Naive cynicism" in everyday theories of responsibility assessment: On biased assumptions of bias. *Journal of Personality and Social Psychology, 76,* 743–753. http://dx.doi.org/10.1037/0022-3514.76.5.743

Leslie, A. M., Knobe, J., & Cohen, A. (2006). Acting intentionally and the side-effect effect. *Psychological Science, 17,* 421–427. http://dx.doi.org/10.1111/j.1467-9280.2006.01722.x

Maguen, S., & Litz, B. (2012). Moral injury in veterans of war. *PTSD Research Quarterly, 23,* 1–6.

Malle, B., & Knobe, J. (1997). The folk concept of intentionality. *Journal of Experimental Social Psychology, 121,* 101–121. http://dx.doi.org/10.1006/jesp.1996.1314

Mazar, N., Amir, O., & Ariely, D. (2008). The dishonesty of honest people: A theory of self-concept maintenance. *Journal of Marketing Research, 45,* 633–644. http://dx.doi.org/10.1509/jmkr.45.6.633

McAdams, D. P. (2011). Narrative identity. In S. J. Schwartz, K. Luyckx, & V. L. Vignoles (Eds.), *Handbook of identity theory and research* (pp. 99–115). New York, NY: Springer. http://dx.doi.org/10.1007/978-1-4419-7988-9_5

McGraw, K. M. (1987). Guilt following transgression: An attribution of responsibility approach. *Journal of Personality and Social Psychology, 53,* 247–256. http://dx.doi.org/10.1037/0022-3514.53.2.247

Meryman, R. (1978). *Mank: The wit, world, and life of Herman Mankiewicz.* New York, NY: Morrow.

Payne, K., & Corrigan, E. (2007). Emotional constraints on intentional forgetting. *Journal of Experimental Social Psychology, 43,* 780–786. http://dx.doi.org/10.1016/j.jesp.2006.07.005

Phelps, E. A., & Sharot, T. (2008). How (and why) emotion enhances the subjective sense of recollection. *Current Directions in Psychological Science, 17,* 147–152. http://dx.doi.org/10.1111/j.1467-8721.2008.00565.x

Pinto-Gouveia, J., & Matos, M. (2011). Can shame memories become a key to identity? The centrality of shame memories predicts psychopathology. *Applied Cognitive Psychology, 25,* 281–290. http://dx.doi.org/10.1002/acp.1689

Pizarro, D. A., Laney, C., Morris, E. K., & Loftus, E. F. (2006). Ripple effects in memory: Judgments of moral blame can distort memory for events. *Memory & Cognition, 34,* 550–555. http://dx.doi.org/10.3758/BF03193578

Pronin, E. (2007). Perception and misperception of bias in human judgment. *Trends in Cognitive Sciences, 11,* 37–43. http://dx.doi.org/10.1016/j.tics.2006.11.001

Pronin, E., Kruger, J., Savitsky, K., & Ross, L. (2001). You don't know me, but I know you: The illusion of asymmetric insight. *Journal of Personality and Social Psychology, 81,* 639–656. http://dx.doi.org/10.1037/0022-3514.81.4.639

Savitsky, K., Medvec, V. H., & Gilovich, T. (1997). Remembering and regretting: The Zeigarnik effect and the cognitive availability of regrettable actions and inactions. *Personality and Social Psychology Bulletin, 23,* 248–257. http://dx.doi.org/10.1177/0146167297233004

Shalvi, S., Gino, F., Barkan, R., & Ayal, S. (2015). Self-serving justifications: Doing wrong and feeling moral. *Current Directions in Psychological Science, 24,* 125–130. http://dx.doi.org/10.1177/0963721414553264

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2013, January). *Life after p-hacking.* Paper presented at the meeting of the Society for Personality and Social Psychology, New Orleans, LA. http://dx.doi.org/10.2139/ssrn.2205186

Stanley, M. L., Henne, P., Iyengar, V., Sinnott-Armstrong, W., & Brigard, F. D. (2017). I'm not the person I used to be: The self and autobiographical memories of immoral actions. *Journal of Experimental Psychology: General, 146,* 884–895. http://dx.doi.org/10.1037/xge0000317

Strohminger, N., & Nichols, S. (2015). Neurodegeneration and identity. *Psychological Science, 26,* 1469–1479. http://dx.doi.org/10.1177/0956797615592381

Tingley, D., Yamamoto, T., Hirose, K., Keele, L., & Imai, K. (2014). Mediation: R package for causal mediation analysis. *Journal of Statistical Software, 59,* 1–38. http://dx.doi.org/10.18637/jss.v059.i05

Waugh, N. C., & Norman, D. A. (1965). Primary memory. *Psychological Review, 72,* 89–104. http://dx.doi.org/10.1037/h0021797

Wohl, M. J. A., & Branscombe, N. R. (2008). Remembering historical victimization: Collective guilt for current ingroup transgressions. *Journal of Personality and Social Psychology, 94,* 988–1006. http://dx.doi.org/10.1037/0022-3514.94.6.988

Zeigarnik, B. (1938). On finished and unfinished tasks. In W. D. Ellis (Ed.), *A source book of Gestalt psychology* (pp. 300–314). London, England: Kegan Paul, Trench, Trubner and Company.

Zitek, E. M., Jordan, A. H., Monin, B., & Leach, F. R. (2010). Victim entitlement to behave selfishly. *Journal of Personality and Social Psychology, 98,* 245–255. http://dx.doi.org/10.1037/a0017168