The phenomenology of remembering our moral transgressions

Shenyang Huang[1] · Matthew L. Stanley[2] · Felipe De Brigard[2,3]

Abstract

People tend to believe that they truly are morally good, and yet they commit moral transgressions with surprising frequency in their everyday lives. To explain this phenomenon, some theorists have suggested that people remember their moral transgressions with fewer details, lower vivacity, and less clarity, relative to their morally good deeds and other kinds of past events. These phenomenological differences are thought to help alleviate psychological discomfort and to help people maintain a morally good self-concept. Given these motivations to alleviate discomfort and to maintain a morally good self-concept, we might expect our more egregious moral transgressions, relative to our more minor transgressions, to be remembered less frequently, with fewer details, with lower vivacity, and with a reduced sense of reliving. More severe moral transgressions might also be less central to constructions of personal identity. In contrast to these expectations, our results suggest that participants' more severe moral transgressions are actually remembered more frequently, more vividly, and with more detail. More severe moral transgressions also tend to be more central to personal identity. We discuss the implications of these results for the motivation to maintain a morally good self-concept and for the functions of autobiographical memory.

Keywords Autobiographical memory · Moral psychology · Identity · Self · Phenomenology

Introduction

Most people believe that they truly are morally good (Aquino & Reed, 2002; Molouki & Bartels, 2017; Strohminger, Knobe, & Newman, 2017), and most people believe that they are superior to both the average person and specific peers in terms of their virtue, compassion, and other positive moral traits (Alicke & Govorun, 2005; Alicke & Sedikides, 2009; Alicke, Vredenburg, Hiatt, & Govorun, 2001; Batson & Collins, 2011; Tappin & McKay, 2017). And yet, people commit moral transgressions with surprising frequency in their everyday lives. In fact, in a recent 3-day study using

✉ Matthew L. Stanley
matthew.stanley@duke.edu

1 Department of Psychology and Neuroscience, Duke University, Durham, NC 27708, USA

2 Department of Psychology and Neuroscience, Center for Cognitive Neuroscience, Duke University, Durham, NC 27708, USA

3 Department of Philosophy, Duke Institute for Brain Sciences, Duke University, Durham, NC 27708, USA

ecological momentary assessment, participants reported 13.6% of the time that they had "committed, were the target of, witnessed, or learned about" an immoral action within the past hour (Hofmann, Wisneski, Brandt, & Skitka, 2014).

If people do commit moral transgressions with such frequency in their everyday lives, why do they continue to adamantly believe that they are morally good? One possible answer to this question involves autobiographical memory, a collection of representations of one's personal past that are of relevance to the self (Conway, 2005; Rubin, 1986). Although psychological research on memory has focused predominantly on passive features of forgetting (e.g., Baddeley, 1990; Schacter, 1999), mounting behavioral and neuroimaging research suggests that there are also active, motivational factors that influence forgetting. That is, people can and do actively undermine their retention of certain unwelcome memories (Anderson & Hanslmayr, 2014; Anderson & Levy, 2009; Bergström, Anderson, Buda, Simons, & Richardson-Klavehn, 2013; Hu, Bergström, Bodenhausen, & Rosenfeld, 2015). Recently, theorists have proposed that people strategically forget their own unethical past behaviors in order to avoid psychological distress and to maintain a morally good self-concept (Kouchaki & Gino, 2016; Stanley & De Brigard, 2019). Specifically, Kouchaki and Gino (2016) suggest that "[people's] memory of their past unethical behavior becomes *less clear, less detailed, and less vivid* over time than their

memory of their ethical actions and of actions unrelated to ethics [emphasis added]"—a phenomenon the authors refer to as *unethical amnesia*.

Kouchaki and Gino (2016) suggest that unethical amnesia impacts both memory *accuracy* (that people are more prone to errors related to memory of their unethical actions) and the *phenomenology* of remembering (that unethical actions, relative to other kinds of memories, are remembered less clearly and vividly). Several studies have examined the former – i.e., memory accuracy – for past unethical actions. For instance, participants who cheated on a task had less accurate memory for the content of an honor code that they had read before the task (Shu, Gino, & Bazerman, 2011). In another study, consumers tended to forget the unethical attributes of the products they purchased (e.g., made from endangered trees or manufactured using child labor); they also regarded such ignorance as more morally permissible than making the purchase while being aware of those unethical attributes (Reczek, Irwin, Zane, & Ehrich, 2018). People do not only forget details of their own past immoral actions to foster their own morally good self-concepts – some evidence suggests that when the rememberer benefits from the unfair actions of others, the rememberer tends to forget critical details about those immoral actions (Bell, Schain, & Echterhoff, 2014).

Although there have been several investigations into memory accuracy for ethical and unethical actions, fewer studies have investigated the *phenomenological characteristics* of remembering unethical actions. In the current two studies, we investigate the relationship between the *moral severity* of past moral transgressions and the phenomenology of remembering them. Moral severity refers to the degree to which a behavior is considered to be more or less morally wrong (Powell & Horne, 2016). For example, even though premeditated murder for no reason and shoplifting are both considered to be morally wrong, premeditated murder for no reason is typically considered to be a much more severe transgression than shoplifting. Our more severe moral transgressions should pose a greater threat to our morally good self-concepts than more minor transgressions. So, memories of more severe moral transgressions should be *more likely* to become obfuscated, relative to more minor transgressions. However, other literature reviewed below might imply the opposite prediction, namely that more severe moral transgressions might be *less likely* to become obfuscated relative to more minor transgressions. In what follows, we investigate three possible effects of moral severity on how people remember and judge their past transgressions, and we attempt to adjudicate between seemingly conflicting predictions implied by different literatures.

First, we investigate a possible moral severity effect on the frequency of recalling moral transgressions. According to Kouchaki and Gino (2016), "people are less likely to think about [memories of unethical actions]"; frequently recalling

our past moral transgressions arouses negative feelings of distress and discomfort and threatens a morally good self-concept. Because memories of more egregious past moral transgressions arouse stronger negative feelings (Escobedo & Adolphs, 2010; Stanley, Henne, Iyengar, Sinnott-Armstrong, & De Brigard, 2017) and pose a greater threat to a morally good self-concept (relative to memories of more minor moral violations), we might expect people to recall their more severe moral transgressions *less frequently* than their more minor transgressions. Some evidence, however, might suggest an alternative possible pattern of results. In a study with offenders of violent crimes, 46% experienced intrusive memories (which are "recurrent … and involuntarily triggered") about their crimes, and 36% reported ruminating on their crimes (Evans, Ehlers, Mezey, & Clark, 2007a, 2007b). These studies suggest that particularly severe moral transgressions might spontaneously and frequently emerge in consciousness with high frequency. However, these studies provide no direct evidence that these convicted criminals actually believed their crimes were morally wrong. So, it remains unclear whether memories of these crimes are damaging to their morally good self-concepts.

Second, we investigate a possible severity effect on the vivacity, detail, and sense of reliving with which past transgressions are recollected. If people selectively obfuscate their memories of their own moral transgressions to mitigate distress and to maintain a belief in a morally good self-concept (Kouchaki & Gino, 2016), we might expect that more severe unethical actions will be the primary targets of obfuscation, because they represent a greater threat to this desired self-concept. Hence, we might expect that more severe moral transgressions, relative to more minor transgressions, will be remembered with less vivacity, fewer sensory details, and a reduced sense of reliving. In contrast to this expectation, however, Evans and collaborators (2007a, 2007b) report that criminals described their intrusive memories of crimes involving harm as "vivid, clear, detailed, intrusive, unchanging, and very distressing." In a related study with criminals convicted of homicide, "murder [was] recalled with a high level of detail, vividness, and sensory elements" as compared to less severe non-homicide offences and positive experiences (Woodworth et al., 2009). These studies suggest that people might recall their severe moral transgressions vividly and with rich detail.

Finally, we investigate a possible moral severity effect on the perceived personal importance and centrality of participants' past moral transgressions. Past experiences remembered as highly central to one's life play significant roles in shaping personal identity, the life story, and self-understanding (Berntsen & Rubin, 2006; Bluck, 2003; Rubin, Schrauf, & Greenberg, 2003). Because people believe they truly are morally good and regard morally good traits as particularly central to their personal identities (e.g., Aquino & Reed, 2002; Fernandez-Duque & Schwartz, 2016; Heiphetz,

Strohminger, & Young, 2017; Strohminger et al., 2017; Strohminger & Nichols, 2014), we might expect people to report that their particularly severe past transgressions are *not* central or important to who they are or who they wish to be. If these severe transgressions were important and central to constructions of personal identity, they might threaten a morally good self-concept. Nevertheless, some evidence suggests that extremely negative events – even those that threaten desired self-concepts – are sometimes particularly central to constructions of personal identity (Rubin, Dennis, & Beckham, 2011; Rubin et al., 2003). A hallmark characteristic of these central negative events is that they tend to be recalled frequently (Rubin et al., 2011). If more severe moral transgressions are, in fact, recalled more frequently than more minor transgressions, they might also be more central to constructions of personal identity.

Study 1

In Study 1, each participant recalled and described six of their own past moral transgressions, and they provided several ratings for each memory. This allowed us to investigate the effects of moral severity on how participants remember and judge their past transgressions.

Method

Participants A total of 138 individuals voluntarily completed this study via Amazon Mechanical Turk (AMT) for monetary compensation. Qualified participants were individuals living in the USA who had completed over 50 HITs with an approval rating above 90%. One participant reported not paying attention (see attention check below) during the study and another participant failed to follow instructions, so data were analyzed with the remaining 136 participants (M_{age} = 33.53 years, SD = 8.51, age range = [21, 66]; 61 females).[1] Informed consent was obtained from each participant in accordance with a protocol approved by the Duke University Campus Institutional Review Board.

Procedure This study consisted of a single, self-paced session. Participants were assured that their names would not be linked to their responses and that all data would be securely saved and de-identified, indefinitely. Participants were asked to recall a total of six events from their personal pasts that occurred when they were older than 15 years of age, and to provide several ratings for each memory. Specifically, participants first recalled and described in two to five sentences one particular event in which they personally did something that they

believed to be at least slightly morally wrong. Participants reported the month and year when the event occurred and rated the severity of the transgression (1 = *slightly morally wrong*; 7 = *very morally wrong*).

Participants then provided ratings on eight different phenomenological characteristics of remembering or memory judgments. All eight items were adapted from the Memory Characteristics Questionnaire (MCQ; Johnson, Foley, Suengas, & Raye, 1988) and the Autobiographical Memory Questionnaire (AMQ; Rubin et al., 2003). Specifically, and directly related to our primary research questions, participants reported their frequency of recalling their moral transgressions, in terms of both voluntary and involuntary retrieval (items 1–2 in Table 1).[2] Participants also rated the vivacity and detail of their memories along with their sense of reliving (items 3–6 in Table 1). Moreover, participants rated the personal importance and centrality of the events (items 7–8 in Table 1). Finally, to address several exploratory questions, participants made ratings on other phenomenological and memory characteristics (see Table S1 in the Supplementary Material for the full list of characteristics). This same process was then repeated for each of the remaining five events, one at a time. After that, participants provided basic demographic information.

At the end, participants were asked the following attention check: "Do you feel that you paid attention, avoided distractions, and took the survey seriously?" They responded by selecting one of the following: (1) no, I was distracted; (2) no, I had trouble paying attention; (3) no, I didn't not take the study seriously; (4) no, something else affected my participation negatively; or (5) yes. Participants were ensured that their responses would not affect their payment or their eligibility for future studies. Only those participants who selected (5) were included in the analyses. This same attention check has been used in recent published work (e.g., Stanley, Yin, & Sinnott-Armstrong, 2019).

Results

Participants described diverse kinds of moral transgressions that varied widely in severity: failing to keep promises to family and friends, damaging the property of other people, infringing upon other people's privacy, shoplifting, cheating on tests and exams, cheating on significant others, shooting voyeuristic films, and hurting animals.

A series of linear mixed-effects models (LMEM) were computed using the 'lme4' software package (Bates, Mächler, Bolker, & Walker, 2015) in R. The judged severity

---

[1] Note that the total number of participants who began the task in each condition but did not submit the HIT is not available.

[2] A memory is recalled *voluntarily* when the rememberer has an intention to deliberately retrieve it prior to actually retrieving it. In contrast, a memory is recalled *involuntarily* when it spontaneously emerges in consciousness without any prior intention to retrieve it (Berntsen, 1996, 2010).

Table 1 Memory characteristic ratings for primary analyses (items 1–8)

Item	1 =	7 =
1. *Voluntary Recall*. Since the event occurred, how often have you willfully thought back to the event or talked about it?	Not at all	Very often
2. *Involuntary Recall*. Since the event occurred, how often has the event popped up in your thoughts by itself, without your attempting to intentionally remember it?	Not at all	Very often
3. *Vivacity*. Overall, how vivid is your memory of the event?	Very vague	Very vivid
4. *Visual Details*. To what extent does this remembered event involve visual details?	No details	Many details
5. *Other Details*. To what extent does this remembered event involve other sensory details (sounds, smells, and/or tastes)?	No details	Many details
6. *Reliving*. While remembering the event, do you feel as though you are mentally reliving it?	Not at all	Completely
7. *Importance*. How important is the event to you personally (it involves an important episode in your life)?	Not at all important	Very important
8. *Centrality*. Is the event in your memory a central part of your life story?	Not at all central	Very central

of the remembered transgression was modeled as the fixed-effect term predicting each memory characteristic in separate models. Participant was modeled as a random effect in each model. Significance for fixed-effects was assessed using Satterthwaite approximations to degrees of freedom, and 95% confidence intervals around beta-values were computed using parametric bootstrapping (500 simulations).

We addressed each of our three main questions in the order that they were presented in the introduction. Table 2 summarizes all LMEM results that address all three main questions (additional exploratory results can be found in Table S2).[3] Table S3 displays the relationships between all memory characteristics. First, we investigated the effect of moral severity on the frequency of voluntary and involuntary retrieval, respectively. Moral transgressions judged to be more severely morally wrong were recalled more frequently – both voluntarily ($b = 0.19$, $p < .0001$, 95% CI [0.13, 0.26]) and involuntarily ($b = 0.23$, $p < .0001$, 95% CI [0.17, 0.29]) – than moral transgressions judged to be less morally wrong.

Second, we investigated the effect of moral severity on reported vivacity, detail, and sense of reliving. Moral transgressions judged to be more severely morally wrong were remembered more vividly ($b = 0.19$, $p < .0001$, 95% CI [0.12, 0.25]) and with more visual ($b = 0.28$, $p < .0001$, 95% CI [0.21, 0.35]) and other sensory details ($b = 0.15$, $p = .0005$, 95% CI [0.08, 0.23]). Remembering more severe moral transgressions also elicited a greater sense of reliving ($b = 0.25$, $p < .0001$, 95% CI [0.19, 0.32]).

Third, we investigated the effect of moral severity on the perceived personal importance and centrality of the moral transgressions. Participants regarded events involving their more severe moral transgressions to be more important ($b =$

0.29, $p < .0001$, 95% CI [0.21, 0.35]) and more central ($b = 0.24$, $p < .0001$, 95% CI [0.17, 0.31]) to their personal identities, as compared to their more minor transgressions.

Because the phenomenology of remembering past events and the judged morality of the actions could differ as a function of how much time has elapsed since the events occurred (Stanley et al., 2017), we statistically controlled for the number of months passed (hereafter referred to as *Months*), starting with the remembered events that occurred in the same month as the experimental session coded as 0 (for a similar methodological approach, see Stanley, Henne, & De Brigard, 2019).[4] An initial LMEM revealed that *Months* had a significant effect on the judged severity of the remembered transgressions ($b = .004$, $p < .0001$, 95% CI [.003, .005]) such that remembered behaviors judged to be more morally wrong occurred in the more distant past. Hence, we statistically controlled for the effect of *Months* by including *Months* as a fixed-effect term in addition to the judged severity of the remembered transgression to predict each memory characteristic in separate models.

The fixed-effect of judged moral severity on all eight characteristic ratings remained significant after statistically controlling for *Months*. Relative to moral transgressions judged to be less morally wrong, moral transgressions judged to be more severely morally wrong were recalled more frequently – both voluntarily ($b = 0.20$, $p < .0001$, 95% CI [0.14, 0.27]) and involuntarily ($b = 0.24$, $p < .0001$, 95% CI [0.19, 0.30]); were remembered more vividly ($b = 0.20$, $p < .0001$, 95% CI [0.12, 0.26]) and with more visual details ($b = 0.27$, $p < .0001$, 95% CI [0.21, 0.33]) and other sensory details ($b = 0.14$, $p = .0005$, 95% CI [0.06, 0.22]); elicited a greater sense of reliving while remembering ($b = 0.25$, $p < .0001$, 95% CI [0.18, 0.31]); and were regarded as more important ($b = 0.29$, $p <$

---

[3] Of note, after applying a Bonferroni correction for multiple comparisons (i.e., adjusted the threshold of significance to .05/21 = .0024), all statistically significant effects identified in Study 1 remained significant.

[4] We thank an anonymous reviewer for this suggestion.

Table 2 Linear mixed-effects models (LMEMs) fixed-effects of judged moral wrongness for Study 1

Outcome variable	b	SE	t	p value	95% CI
Voluntary Recall	0.19	0.03	5.69	< .0001	[0.13, 0.26]
Involuntary Recall	0.23	0.03	7.11	< .0001	[0.17, 0.29]
Vivacity	0.19	0.03	5.94	< .0001	[0.12, 0.25]
Visual Details	0.28	0.03	8.46	< .0001	[0.21, 0.35]
Other Details	0.15	0.04	3.94	< .0001	[0.08, 0.23]
Reliving	0.25	0.03	7.65	< .0001	[0.19, 0.32]
Importance	0.29	0.04	7.81	< .0001	[0.21, 0.35]
Centrality	0.24	0.04	6.86	< .0001	[0.17, 0.31]

SE standard error, *CI* confidence interval

.0001, 95% CI [0.21, 0.37]) and more central ($b = 0.24$, $p <$ .0001, 95% CI [0.17, 0.31]) to their personal identities.

Discussion

The results from Study 1 suggest that, even though more severe moral transgressions presumably pose a greater threat to morally good self-concepts than more minor transgressions, more severe moral transgressions were recalled more frequently, with more details, with greater vivacity, and with a heightened sense of reliving. More severe moral transgressions were also more important and central to participants' personal identities.

Study 2

In order to more directly compare the phenomenology of remembering transgressions of varying degrees of moral severity, in Study 2 we explicitly cued participants to provide a memory of a slightly morally wrong past transgression and a very morally wrong past transgression (or the most morally wrong transgression that they had committed). We then directly compared phenomenological characteristics and memory judgments as a function of these two cues.

Method

Participants A total of 171 individuals voluntarily completed this study via AMT for monetary compensation. Qualified participants were individuals living in the USA who had completed over 50 HITs with an approval rating above 90%. Six participants reported not paying attention during the study (see attention check below), and three other participants failed to follow instructions, so data were analyzed with the remaining 163 participants ($M_{age} = 34.20$ years, $SD = 9.61$, age range = [19, 69]; 72 females). Informed consent was obtained from

each participant in accordance with a protocol approved by the Duke University Campus Institutional Review Board.

Procedure This study consisted of a single, self-paced session. Participants were assured that their names would not be linked to their responses and that all data would be securely saved and de-identified, indefinitely. Participants were asked to recall two specific events that happened when they were older than 15 years of age and to provide ratings for both events. Specifically, participants recalled and described in two to five sentences an event in which they personally did something that they believe to be *slightly morally wrong*, and an event in which they personally did something they believe to be *very morally wrong* (or the most morally wrong thing they have done since they were 15 years old). For each remembered event, participants indicated the month and year that the event occurred and provided a moral judgment (1 = *very morally wrong*; 7 = *very morally right*). Participants also provided ratings on the exact same items as in Study 1 (see Table 1 and Table S1). Then, participants reported basic demographic information.

Finally, as in Study 1, participants were asked whether they paid attention, avoided distractions, and took the survey seriously. They responded by selecting one of the following: (1) no, I was distracted; (2) no, I had trouble paying attention; (3) no, I didn't not take the study seriously; (4) no, something else affected my participation negatively; or (5) yes. Participants were ensured that their responses would not affect their payment or their eligibility for future studies. Only those participants who selected (5) were included in the analyses.

Results

Participants described diverse kinds of moral transgressions in response to both memory cues. Slightly morally wrong behaviors included making graffiti, not returning extra change, shoplifting small items, cheating on exams, and lying to family and friends for selfish reasons. For very morally wrong behaviors, participants reported cheating on significant others, inflicting physical harm on others, drunk driving, and using drugs, among other things.

In order to confirm that our memory cue manipulation worked as expected, we conducted an initial paired-samples *t*-test. This indicated that *slightly morally wrong*-cued memories ($M_{slightly} = 3.04$, $SD = .76$) were judged to be less morally wrong than *very morally wrong*-cued memories ($M_{very} = 1.94$, $SD = 1.59$; $t(162) = 8.77$, $p < .0001$, 95% CI [0.85, 1.35], Cohen's $d = .69$).

We then computed paired samples *t*-tests to compare each memory characteristic as a function of *Memory Cue* (i.e., *slightly morally wrong* versus *very morally wrong*). Table 3 summarizes the paired samples *t*-tests results addressing our

Table 3 Paired-samples *t*-test results for differences in memory characteristics as a function of cue (slightly vs. very morally wrong) for Study 2 (items 1–8)

Outcome variable	Mean (SD)		$t(162)$	p value	Cohen's d
	slightly	very			
Voluntary Recall	2.83 (1.75)	3.71 (1.81)	4.97	< .0001	0.39
Involuntary Recall	2.72 (1.57)	4.16 (1.67)	9.11	< .0001	0.71
Vivacity	5.24 (1.51)	5.65 (1.38)	2.84	.0050	0.22
Visual Details	4.46 (1.64)	5.37 (1.41)	6.59	< .0001	0.52
Other Details	3.18 (1.76)	4.03 (1.90)	5.78	< .0001	0.45
Reliving	4.74 (1.62)	5.29 (1.53)	4.01	< .0001	0.31
Importance	2.84 (1.78)	4.36 (1.83)	8.11	< .0001	0.64
Centrality	2.37 (1.66)	3.70 (1.86)	7.66	< .0001	0.60

three main questions that were presented in the Introduction (additional exploratory results can be found in Table S4). (Table S5 and Table S6 provide correlation matrices for all relationships between all measured memory characteristics.) First, we investigated the effect of moral severity on the frequency of voluntary and involuntary retrieval, respectively. Past behaviors judged to be very morally wrong (i.e., severe moral transgressions) were more frequently recalled – both voluntarily ($t(162) = 4.97$, $p < .0001$, Cohen's $d = 0.39$) and involuntarily ($t(162) = 9.11$, $p < .0001$, Cohen's $d = 0.71$) – relative to behaviors judged to be slightly morally wrong (i.e., minor moral transgressions). Figure 1 graphically depicts the differences in frequency of recall as a function of the severity of the moral transgression.

Second, we investigated the effect of moral severity on the judged vivacity, detail, and sense of reliving. Severe moral transgressions were remembered more vividly ($t(162) = 2.84$, $p = .0050$, Cohen's $d = 0.22$) and with more visual
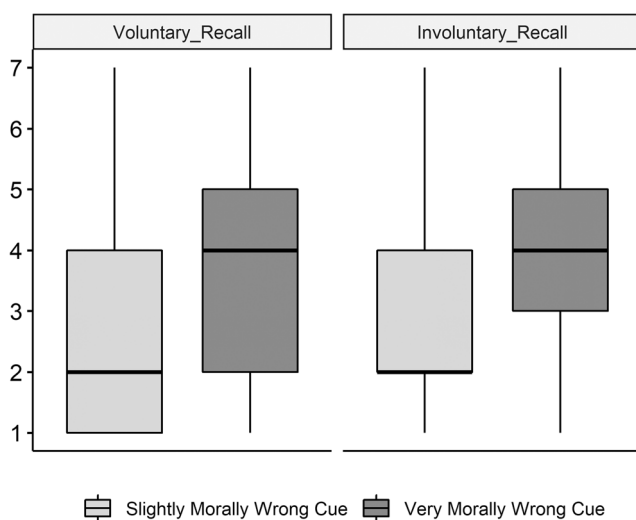

Fig. 1 Box plots are depicted for the frequency of voluntary recall and involuntary recall as a function of memory cue (*slightly morally wrong* vs. *very morally wrong*)

details ($t(162) = 6.59$, $p < .0001$, Cohen's $d = 0.52$) and other sensory details ($t(162) = 5.78$, $p < .0001$, Cohen's $d = 0.45$) than minor moral transgressions. Remembering severe moral transgressions also elicited a greater sense of reliving ($t(162) = 4.01$, $p < .0001$, Cohen's $d = 0.31$). See Fig. 2 for differences in *Vivacity*, *Visual Details*, *Other Details*, and *Reliving* as functions of moral severity.

Third, we investigated the effect of moral severity on the perceived personal importance and centrality of their past moral transgressions. Participants considered their severe moral transgressions to be more personally important $t(162) = 8.11$, $p < .0001$, Cohen's $d = 0.64$) and more central ($t(162) = 7.66$, $p < .0001$, Cohen's $d = 0.60$) to their personal identities than their minor moral transgressions. See Fig. 3.[5]

As in Study 1, we statistically controlled for *Months* (the number of months since the event occurred) in Study 2. A paired samples *t*-test revealed that *slightly morally wrong*-cued memories ($M_{slightly} = 55.13$, $SD = 81.32$) occurred in more recent past than *very morally wrong*-cued memories ($M_{very} = 106.89$, $SD = 111.09$; $t(162) = 5.56$, $p < .0001$, 95% CI [33.39, 70.13], Cohen's $d = 0.44$). Hence, we statistically controlled for the effect of *Months* by running a series of ANCOVAs with *Memory Cue* (i.e., *slightly morally wrong* vs. *very morally wrong*) as the factor and *Months* as the covariate for each memory characteristic rating. The effects of Memory Cue on all eight characteristic ratings remained significant after statistically controlling for *Months*. Compared to minor moral transgressions, severe moral transgressions were more frequently recalled – both voluntarily ($F(1, 324) = 21.1$, $p < .0001$, $\eta^2 = .06$) and involuntarily ($F(1, 324) = 65.4$, $p < .0001$, $\eta^2 = .17$); were remembered more vividly ($F(1, 324) = 7.3$, $p = .0074$, $\eta^2 = .02$) and with more visual details ($F(1,

---

[5] Of note, after applying a Bonferroni correction for multiple comparisons (i.e., adjusted the threshold of significance to .05/21 = .0024), all significant effects in Study 2 remained significant, except for one (the *vivacity* of the memories of transgressions of different moral severity)
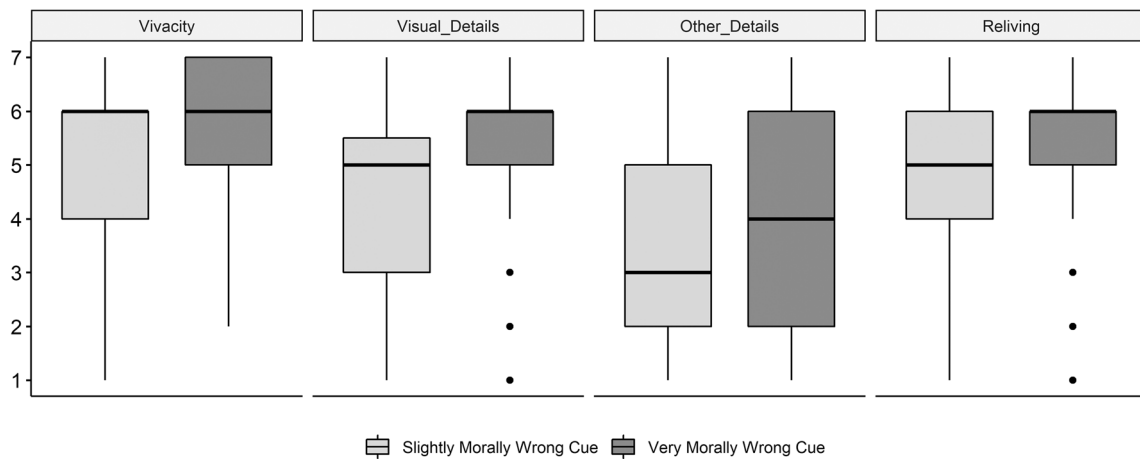
Fig. 2 Box plots are depicted for the vivacity, detail, and sense of reliving as a function of memory cue (*slightly morally wrong* vs. *very morally wrong*)

324) = 25.6, p < .0001, η^2 = .07) and other sensory details (F(1, 324) = 15.6, p < .0001, η^2 = .05); elicited a greater sense of reliving (F(1, 324) = 8.5, p = .0038, η^2 = .03) while remembering; were regarded as more personally important (F(1, 324) = 51.8, p < .0001, η^2 = .14) and more central (F(1, 324) = 40.2, p < .0001, η^2 = .11) to their personal identities.[6]

We also performed more conservative non-parametric Wilcoxon signed-rank tests for paired samples (see Table S4). We obtained the same pattern of results with parametric and non-parametric tests (Figs. 1 and 2 and Fig. S1).

Discussion

These findings from Study 2 corroborate our results from Study 1 using a different cueing procedure. In Study 1, LMEMs revealed that remembering more severe moral transgressions, relative to less severe moral transgressions, is associated with higher ratings on memory characteristics. In Study 2, when we explicitly cued participants to provide two different kinds of memories differing in terms of the judged severity of the transgressions (*slightly morally wrong* vs. *very morally wrong*), we obtained the same pattern of results for the memory characteristics of interest.

General discussion

In an influential recent study, Kouchaki and Gino (2016) found evidence that people's memories of their own past unethical behaviors become obfuscated over time; they refer to this phenomenon as *unethical amnesia*. The authors offer an explanation for this unethical amnesia phenomenon: unethical

---

[6] Of note, after applying a Bonferroni correction for multiple comparisons (i.e., adjusted the threshold of significance to .05/21 = .0024), all significant effects in Study 2 remained significant, except for two (the *vivacity* of the memories of transgressions of different moral severity and the sense of *reliving* through the memory).
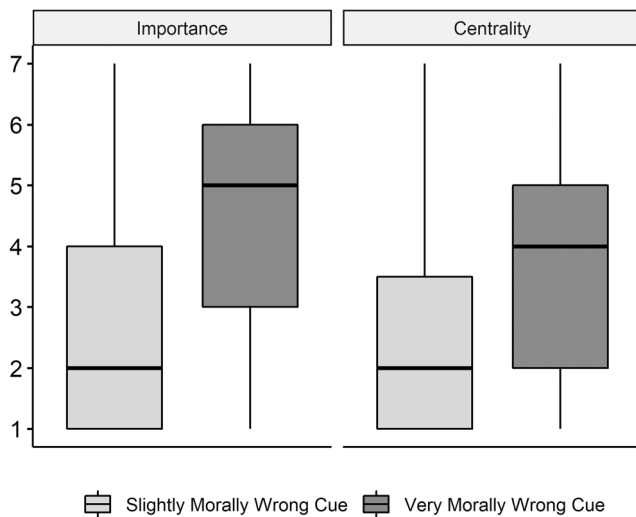
past behaviors become obfuscated so that people can avoid psychological distress and protect against threats to their morally good self-concept. As we have argued, more severe moral transgressions elicit stronger negative feelings (Escobedo & Adolphs, 2010; Stanley et al., 2017) and should pose a greater threat to our morally good self-concepts (relative to more minor transgressions). For this reason, memories of more severe moral transgressions should be *more likely* to become obfuscated relative to more minor transgressions. However, our results instead suggest that recollecting more severe moral transgressions, relative to less severe moral transgressions, occurred more frequently, was experienced with greater vivacity, detail, and sense of reliving, and was judged to be more important and central to constructions of personal identity. Our findings are more consistent with studies investigating how convicted criminals remember their crimes (e.g., Evans et al., 2007a, 2007b; Woodworth et al., 2009), according to which convicted criminals tend to frequently ruminate on their crimes, their crimes tend to be remembered vividly, and the experience of remembering them tends to be accompanied with strong negative emotions.

The frequent retrieval, heightened vivacity, and substantial detail of memories involving our more severe moral transgressions are likely threatening to our morally good self-concepts. Nevertheless, the vast majority of people do still believe that they are fundamentally and essentially morally good (Aquino & Reed, 2002; Strohminger et al., 2017). They believe that they exemplify desirable moral traits, qualities, and virtues (Aquino & Reed, 2002; Stanley & De Brigard, 2019) and that these traits, qualities, and virtues are essential to their constructions of their selves and identities (Molouki & Bartels, 2017; Strohminger et al., 2017). How is it, then, that so many people hold these beliefs despite readily and frequently recalling their severe moral transgressions and remembering them vividly and with considerable detail? We have suggested that other mechanisms involving autobiographical memory might still help people to support their beliefs in their morally

Fig. 3 Box plots are depicted for the perceived personal importance and centrality of past moral transgressions as a function of memory cue (*slightly morally wrong* vs. *very morally wrong*)

good self-concepts (Stanley & De Brigard, 2019). For example, people strategically compare their current selves to their past selves in order to facilitate an impression of moral improvement over time (Stanley & De Brigard, 2019; Stanley et al., 2019, 2017). So, frequently and vividly remembering our severe past transgressions might help us to continually *contrast* our more morally upstanding current selves with our less morally upstanding past selves. In this way, we can come to believe that we have changed for the better since committing those transgressions, and that who we are now is different than who we were then.

Readily, frequently, and vividly remembering our particularly severe past transgressions with a high degree of detail might also serve other functions. For example, memories of specific past events often serve a *directive function* (Pillemer, 2000, 2003). That is, our memories often serve as guides and reference points for forming intentions and making decisions. In one study, for example, Mackey et al. (2001) provided a detailed report of adolescent's memory involving a particular interaction with her parents. This remembered event came to mind frequently and continued to direct social problem solving in many new circumstances after the initial occurrence. Similarly, remembering our particularly severe past transgressions readily, frequently, and vividly might be helpful for forming of intentions to behave differently in the future and for actually making better decisions in the future – even if those severe transgressions tend to be especially unpleasant and distressing to remember and ruminate upon.

A notable feature of our studies is that we utilized real-world moral transgressions that actually occurred in the past. In contrast, a relatively large proportion of studies in moral psychology have relied on artificial vignettes describing fictional dilemmas (e.g., the trolley problem or other sacrificial dilemmas) to probe people's moral judgments and decisions.

However, the kinds of moral transgressions that people actually commit in the world rarely match the kinds of transgressions that are described in the vignettes frequently used in the literature (Hofmann et al., 2014). Furthermore, recent empirical evidence (e.g., Bostyn, Sevenhant, & Roets, 2018; FeldmanHall et al., 2012) even suggests that moral judgments and decisions made in hypothetical vignettes fail to accurately predict similar kinds of real-life moral behaviors. However, there has been some debate over the conclusiveness of evidence provided by Bostyn et al. (2018) and over the value of research using hypothetical vignettes (Białek, Turpin, & Fugelsang, 2019; Bostyn & Roets, 2019; Evans & Brandt, 2019; Plunkett & Greene, 2019). In our studies, we obtained participants' autobiographical memories of specific past moral transgressions to gain insights into how they remember and think about those events.

Several studies have investigated the *objective accuracy* of memories involving past moral and immoral actions (e.g., Reczek et al., 2018; Shu et al., 2011). However, the phenomenology of remembering these actions and the judgments people make about them have remained largely unexplored (Stanley, Yang, & De Brigard, 2018). This is surprising given considerable evidence for dissociations between phenomenological characteristics (e.g., vivacity and detail) and objective accuracy (Berntsen & Thomsen, 2005; Roediger & McDermott, 1995; Talarico & Rubin, 2003, 2007). Our studies provide some insights into the phenomenology of remembering past moral transgressions.

Limitations and conclusions

One limitation of our approach is that, because we did not experimentally manipulate the severity of the moral transgressions, we could only obtain a correlative understanding of relationships between judged moral severity and memory characteristics. This entails that we could not draw strong conclusions about the direction of our effects, nor could we examine whether third variables are casually responsible for other identified bivariate relationships. We also relied on self-reports, as is the case for most research on autobiographical memories for real-life events, which has certain inherent limitations (see Garcia & Gustavson, 1997). Furthermore, it is possible that participants found certain questions ambivalent. For instance, in responding to the *Importance* and the *Centrality* questions (i.e., "How important is the event to you personally (it involves an important episode in your life)?" and "Is the event in your memory a central part of your life story?"), participants might view a moral transgression to be important and central to one's life story *either* because one thinks the event is indicative of one's current identity *or* because one has learned from the event and since become a better person.

Our findings leave also open the possibility that only *a few* of people's moral transgressions are frequently and vividly remembered while most memories of past immoral behaviors become obfuscated with time. Given the relatively limited quantities of memories our participants reported (six memories in Study 1 and two memories in Study 2), our results do not rule out this possibility. In order to test the proportion of clearly remembered memories of moral transgressions, future work may need to collect a larger quantity of autobiographical memories of moral transgressions per participant. To that end, researchers may find that showing the participants common words as cues (see Escobedo & Adolphs (2010) for a similar approach), a variant of the Galton cue-word technique (Rubin & Schulkind, 1997). Alternatively, researchers could ask participants to provide more specific kinds of moral transgressions (e.g., lying, causing emotional harm; see Stanley et al. (2017) for a similar approach) to facilitate autobiographical memory retrieval.

Open Practices Statement The data and materials for the experiments reported here are available upon request to the corresponding author, and none of the experiments were preregistered.

Compliance with ethical standards

Declaration of conflicting interests The authors declare that they have no conflicts of interest with respect to their authorship or the publication of this article.

References

Alicke, M. D., & Govorun, O. (2005). The Better-Than-Average Effect. In *Studies in Self and Identity. The Self in Social Judgment* (pp. 85–106). New York, NY, US: Psychology Press.

Alicke, M. D., & Sedikides, C. (2009). Self-enhancement and self-protection: What they are and what they do. *European Review of Social Psychology*, *20*(1), 1–48. doi:https://doi.org/10.1080/10463280802613866

Alicke, M. D., Vredenburg, D. S., Hiatt, M., & Govorun, O. (2001). The "Better Than Myself Effect." *Motivation and Emotion*, *25*(1), 7–22. doi:https://doi.org/10.1023/A:1010655705069

Anderson, M. C., & Hanslmayr, S. (2014). Neural mechanisms of motivated forgetting. *Trends in Cognitive Sciences*, *18*(6), 279–292. doi: https://doi.org/10.1016/j.tics.2014.03.002

Anderson, M. C., & Levy, B. J. (2009). Suppressing unwanted memories. *Current Directions in Psychological Science*, *18*(4), 189–194. doi: https://doi.org/10.1111/j.1467-8721.2009.01634.x

Aquino, K., & Reed, A. I. (2002). The self-importance of moral identity. *Journal of Personality and Social Psychology*, *83*(6), 1423–1440. doi:https://doi.org/10.1037/0022-3514.83.6.1423

Baddeley, A. (1990). *Human memory: Theory and practice*. Retrieved from https://find.library.duke.edu/catalog/DUKE000893041

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*(1), 1–48.

Batson, C. D., & Collins, E. C. (2011). Moral hypocrisy: A self-enhancement/self-protection motive in the moral domain. In Handbook of self-enhancement and self-protection (pp. 92–111). New York, NY, US: The Guilford Press.

Bell, R., Schain, C., & Echterhoff, G. (2014). How selfish is memory for cheaters? Evidence for moral and egoistic biases. *Cognition*, *132*(3), 437–442. doi:https://doi.org/10.1016/j.cognition.2014.05.001

Bergström, Z. M., Anderson, M. C., Buda, M., Simons, J. S., & Richardson-Klavehn, A. (2013). Intentional retrieval suppression can conceal guilty knowledge in ERP memory detection tests. *Biological Psychology*, *94*(1), 1–11. doi:https://doi.org/10.1016/j.biopsycho.2013.04.012

Berntsen, D. (1996). Involuntary autobiographical memories. *Applied Cognitive Psychology*, *10*(5), 435–454. doi:https://doi.org/10.1002/(SICI)1099-0720(199610)10:5<435::AID-ACP408>3.0.CO;2-L

Berntsen, D. (2010). The unbidden past: Involuntary autobiographical memories as a basic mode of remembering. *Current Directions in Psychological Science*, *19*(3), 138–142. doi:https://doi.org/10.1177/0963721410370301

Berntsen, D., & Rubin, D. C. (2006). The centrality of event scale: A measure of integrating a trauma into one's identity and its relation to post-traumatic stress disorder symptoms. *Behaviour Research and Therapy*, *44*(2), 219–231. doi:https://doi.org/10.1016/j.brat.2005.01.009

Berntsen, D., & Thomsen, D. K. (2005). Personal memories for remote historical events: Accuracy and clarity of flashbulb memories related to World War II. *Journal of Experimental Psychology General*, *134*(2), 242–257. doi:https://doi.org/10.1037/0096-3445.134.2.242

Białek, M., Turpin, M. H., & Fugelsang, J. A. (2019). What Is the Right Question for Moral Psychology to Answer? Commentary on Bostyn, Sevenhant, and Roets (2018). *Psychological Science*, *30*(9), 1383–1385. doi:https://doi.org/10.1177/0956797618815171

Bluck, S. (2003). Autobiographical memory: Exploring its functions in everyday life. *Memory*, *11*(2), 113–123. doi:https://doi.org/10.1080/741938206

Bostyn, D. H., & Roets, A. (2019). Should Trolleys Be Scared of Mice? Replies to Evans and Brandt (2019); Białek, Turpin, and Fugelsang (2019); Colman, Gold, and Pulford (2019); and Plunkett and Greene (2019). *Psychological Science*, *30*(9), 1392–1396. doi:https://doi.org/10.1177/0956797619865236

Bostyn, D. H., Sevenhant, S., & Roets, A. (2018). Of Mice, Men, and Trolleys: Hypothetical Judgment Versus Real-Life Behavior in Trolley-Style Moral Dilemmas. *Psychological Science*, *29*(7), 1084–1093. doi:https://doi.org/10.1177/0956797617752640

Conway, M. A. (2005). Memory and the self. *Journal of Memory and Language*, *53*(4), 594–628. doi:https://doi.org/10.1016/j.jml.2005.08.005

Escobedo, J. R., & Adolphs, R. (2010). Becoming a better person: Temporal remoteness biases autobiographical memories for moral events. *Emotion*, *10*(4), 511–518. doi:https://doi.org/10.1037/a0018723

Evans, A. M., & Brandt, M. J. (2019). Comparing the Effects of Hypothetical Moral Preferences on Real-Life and Hypothetical Behavior: Commentary on Bostyn, Sevenhant, and Roets (2018). *Psychological Science*, *30*(9), 1380–1382. doi:https://doi.org/10.1177/0956797618815482

Evans, C., Ehlers, A., Mezey, G., & Clark, D. M. (2007a). Intrusive memories and ruminations related to violent crime among young offenders: Phenomenological characteristics. *Journal of Traumatic Stress*, *20*(2), 183–196. doi:https://doi.org/10.1002/jts.20204

Evans, C., Ehlers, A., Mezey, G., & Clark, D. M. (2007b). Intrusive memories in perpetrators of violent crime: Emotions and cognitions. *Journal of Consulting and Clinical Psychology*, *75*(1), 134–144. doi:https://doi.org/10.1037/0022-006X.75.1.134

FeldmanHall, O., Mobbs, D., Evans, D., Hiscox, L., Navrady, L., & Dalgleish, T. (2012). What we say and what we do: The relationship

between real and hypothetical moral choices. *Cognition*, *123*(3), 434–441. doi:https://doi.org/10.1016/j.cognition.2012.02.001

Fernandez-Duque, D., & Schwartz, B. (2016). Common Sense Beliefs about the Central Self, Moral Character, and the Brain. *Frontiers in Psychology*, *6*. doi:https://doi.org/10.3389/fpsyg.2015.02007

Garcia, J., & Gustavson, A. R. (1997). The Science of Self-Report. *APS Observer*, *10*(1). Retrieved from https://www.psychologicalscience.org/observer/the-science-of-self-report

Heiphetz, L., Strohminger, N., & Young, L. L. (2017). The Role of Moral Beliefs, Memories, and Preferences in Representations of Identity. *Cognitive Science*, *41*(3), 744–767. doi:https://doi.org/10.1111/cogs.12354

Hofmann, W., Wisneski, D. C., Brandt, M. J., & Skitka, L. J. (2014). Morality in everyday life. *Science; Washington*, *345*(6202), 1340–1343. http://dx.doi.org.proxy.lib.duke.edu/10.1126/science.1251560

Hu, X., Bergström, Z. M., Bodenhausen, G. V., & Rosenfeld, J. P. (2015). Suppressing Unwanted Autobiographical Memories Reduces Their Automatic Influences: Evidence From Electrophysiology and an Implicit Autobiographical Memory Test. *Psychological Science*, *26*(7), 1098–1106. doi:https://doi.org/10.1177/0956797615575734

Johnson, M. K., Foley, M. A., Suengas, A. G., & Raye, C. L. (1988). Phenomenal characteristics of memories for perceived and imagined autobiographical events. *Journal of Experimental Psychology: General*, *117*(4), 371–376. doi:https://doi.org/10.1037/0096-3445.117.4.371

Kouchaki, M., & Gino, F. (2016). Memories of unethical actions become obfuscated over time. *Proceedings of the National Academy of Sciences*, *113*(22), 6166–6171. doi:https://doi.org/10.1073/pnas.1523586113

Mackey, K., Arnold, M. L., & Pratt, M. W. (2001). Adolescents' Stories of Decision Making in More and Less Authoritative Families: Representing the Voices of Parents in Narrative. *Journal of Adolescent Research*, *16*(3), 243–268. doi:https://doi.org/10.1177/0743558401163001

Molouki, S., & Bartels, D. M. (2017). Personal change and the continuity of the self. *Cognitive Psychology*, *93*, 1–17. doi:https://doi.org/10.1016/j.cogpsych.2016.11.006

Pillemer, D. B. (2000). *Momentous Events, Vivid Memories*. Retrieved from http://ebookcentral.proquest.com/lib/duke/detail.action?docID=3300311

Pillemer, D. B. (2003). Directive functions of autobiographical memory: The guiding power of the specific episode. *Memory*, *11*(2), 193–202. doi:https://doi.org/10.1080/741938208

Plunkett, D., & Greene, J. D. (2019). Overlooked Evidence and a Misunderstanding of What Trolley Dilemmas Do Best: Commentary on Bostyn, Sevenhant, and Roets (2018). *Psychological Science*, *30*(9), 1389–1391. doi:https://doi.org/10.1177/0956797619827914

Powell, D., & Horne, Z. (2016). *Moral severity is represented as a domain-general magnitude*. doi:None

Reczek, R. W., Irwin, J. R., Zane, D. M., & Ehrich, K. R. (2018). That's Not How I Remember It: Willfully Ignorant Memory for Ethical Product Attribute Information. *Journal of Consumer Research*, *45*(1), 185–207. doi:https://doi.org/10.1093/jcr/ucx120

Roediger, H. L., & McDermott, K. B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*(4), 803–814. doi:https://doi.org/10.1037/0278-7393.21.4.803

Rubin, D. C. (1986). *Autobiographical memory*. Retrieved from https://find.library.duke.edu/catalog/DUKE000660717

Rubin, D. C., Dennis, M. F., & Beckham, J. C. (2011). Autobiographical memory for stressful events: The role of autobiographical memory in posttraumatic stress disorder. *Consciousness and Cognition*, *20*(3), 840–856. doi:https://doi.org/10.1016/j.concog.2011.03.015

Rubin, D. C., Schrauf, R. W., & Greenberg, D. L. (2003). Belief and recollection of autobiographical memories. *Memory & Cognition*, *31*(6), 887–901. doi:https://doi.org/10.3758/BF03196443

Rubin, D. C. & Schulkind, M. D. (1997). Distribution of important and word-cued autobiographical memories in 20-, 35-, and 75- year-old adults. *Psychology and Aging, 12*, 524-535.

Schacter, D. L. (1999). The seven sins of memory. Insights from psychology and cognitive neuroscience. *The American Psychologist*, *54*(3), 182–203.

Shu, L. L., Gino, F., & Bazerman, M. H. (2011). Dishonest Deed, Clear Conscience: When Cheating Leads to Moral Disengagement and Motivated Forgetting. *Personality and Social Psychology Bulletin*, *37*(3), 330–349. doi:https://doi.org/10.1177/0146167211398138

Stanley, M. L., & De Brigard, F. (2019). Moral memories and the belief in the good self. *Current Directions in Psychological Science*, 0963721419847990. doi:https://doi.org/10.1177/0963721419847990

Stanley, M. L., Henne, P., & De Brigard, F. (2019). Remembering moral and immoral actions in constructing the self. *Memory & Cognition*, *47*(3), 441–454. doi:https://doi.org/10.3758/s13421-018-0880-y

Stanley, M. L., Henne, P., Iyengar, V., Sinnott-Armstrong, W., & De Brigard, F. (2017). I'm not the person I used to be: The self and autobiographical memories of immoral actions. *Journal of Experimental Psychology: General*, *146*(6), 884–895. doi:https://doi.org/10.1037/xge0000317

Stanley, M. L., Yang, B. W., & De Brigard, F. (2018). No evidence for unethical amnesia for imagined actions: A failed replication and extension. *Memory & Cognition*, *46*(5), 787–795. doi:https://doi.org/10.3758/s13421-018-0803-y

Strohminger, N., Knobe, J., & Newman, G. (2017). The true self: A psychological concept distinct from the self. *Perspectives on Psychological Science*, *12*(4), 551–560. doi:https://doi.org/10.1177/1745691616689495

Strohminger, N., & Nichols, S. (2014). The essential moral self. *Cognition*, *131*(1), 159–171. doi:https://doi.org/10.1016/j.cognition.2013.12.005

Talarico, J. M., & Rubin, D. C. (2003). Confidence, not consistency, characterizes flashbulb memories. *Psychological Science*, *14*(5), 455–461. doi:https://doi.org/10.1111/1467-9280.02453

Talarico, J. M., & Rubin, D. C. (2007). Flashbulb memories are special after all; in phenomenology, not accuracy. *Applied Cognitive Psychology*, *21*(5), 557–578. doi:https://doi.org/10.1002/acp.1293

Tappin, B. M., & McKay, R. T. (2017). The illusion of moral superiority. *Social Psychological and Personality Science*, *8*(6), 623–631. doi:https://doi.org/10.1177/1948550616673878

Woodworth, M., Porter, S., ten Brinke, L., Doucette, N. L., Peace, K., & Campbell, M. A. (2009). A comparison of memory for homicide, non-homicidal violence, and positive life experiences. *International Journal of Law and Psychiatry*, *32*(5), 329–334. doi:https://doi.org/10.1016/j.ijlp.2009.06.008