



# Making moral principles suit yourself

Matthew L. Stanley<sup>1</sup> · Paul Henne<sup>2</sup> · Laura Niemi<sup>3</sup> · Walter Sinnott-Armstrong<sup>4</sup> · Felipe De Brigard<sup>5</sup>

Accepted: 12 April 2021  
© The Psychonomic Society, Inc. 2021

## Abstract

Normative ethical theories and religious traditions offer general moral principles for people to follow. These moral principles are typically meant to be fixed and rigid, offering reliable guides for moral judgment and decision-making. In two preregistered studies, we found consistent evidence that agreement with general moral principles shifted depending upon events recently accessed in memory. After recalling their own personal violations of moral principles, participants agreed less strongly with those very principles—relative to participants who recalled events in which other people violated the principles. This shift in agreement was explained, in part, by people’s willingness to excuse their own moral transgressions, but not the transgressions of others. These results have important implications for understanding the roles memory and personal identity in moral judgment. People’s commitment to moral principles may be maintained when they recall others’ past violations, but their commitment may wane when they recall their own violations.

**Keywords** Morality · Principles · Memory · Autobiographical · Identity

Most normative ethical theories and religious traditions offer general moral principles that dictate how people ought to act across different situations. For example, Kantian deontologists argue that people should always respect the dignity of all persons and always treat people as an end, never merely as a means (Kant, 1785/1959). Some utilitarians argue that people should always act in a way that achieves the greatest good for the greatest number (Mill, 1861/1998). The Ten Commandments in Christian religious traditions offer a set of moral principles for the faithful to follow. In making everyday judgments and decisions, people appeal to a variety of different moral principles (Albertzart, 2013; Väyrynen, 2008): One might assert that people should always be honest, treat others fairly, show respect to authority figures, or help others in need. Such moral principles are meant to be stable

guides for making future decisions and for justifying evaluations of past actions. That is, they are meant to be fixed and insensitive to any individual’s immediate circumstances or needs.

But do moral principles stably guide and justify people’s actions and judgments? Some cognitive science research suggests that people may use general moral principles to make and change their judgments and decisions with some reliability (Cushman et al., 2006; Horne et al., 2015). For example, people commonly appeal to the *action principle* (i.e., that it is worse to cause harm by direct action than to cause equivalent harms by omission) to justify their moral judgments across diverse scenarios (Cushman & Young, 2011; Cushman et al., 2006). Furthermore, whether people agree with a utilitarian moral principle reliably predicts the judgments they

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

<sup>1</sup> Department of Psychology and Neuroscience, Center for Cognitive Neuroscience, Duke University, 417 Chapel Drive, Durham, NC 27708, USA

<sup>2</sup> Department of Philosophy, Neuroscience Program, Lake Forest College, Lake Forest, IL, USA

<sup>3</sup> Department of Psychology, College of Arts and Sciences; Management and Organizations, Charles H. Dyson School of Applied Economics and Management, Cornell SC Johnson College of Business, Cornell University, Ithaca, NY, USA

<sup>4</sup> Department of Philosophy, Kenan Institute for Ethics, Center for Cognitive Neuroscience, Duke Institute for Brain Sciences, Law School, Duke University, Durham, NC, USA

<sup>5</sup> Department of Philosophy, Department of Psychology and Neuroscience, Center for Cognitive Neuroscience, Duke Institute for Brain Sciences, Duke University, Durham, NC 27708, USA

make in moral dilemmas (Lombrozo, 2009), and reminding people of an endorsed utilitarian principle induces judgment change in moral dilemmas in a way that accords with the principle (Home et al., 2015). More anecdotally, many of us can recall situations in which we chose not to perform an action that would benefit us in order to abide by a particular moral principle, such as “Keep your promises.” Despite some evidence that people use moral principles in making and justifying their judgments and decisions, they still might flexibly shift their agreement with moral principles depending upon the events that are accessed in memory. In particular, we suggest that people might hold moral principles with less conviction after recalling their own violations of those principles yet maintain their level of conviction after recalling others’ violations of those same principles.

Why might people reduce agreement with moral principles after recollecting their own violations of those principles, but not the violations of others? One possible answer to this question is that people tend to *excuse* their own moral transgressions, but not the moral transgressions of others. Selectively excusing their own past transgressions more than the transgressions of others may motivate them to reduce the extent to which they agree with the violated moral principles. That is, excusing our own improprieties may highlight that the principle should be held less rigidly or with less conviction. Lending credence to this possible explanation, converging lines of research suggest that people evaluate themselves more favorably than others (Alicke & Sedikides, 2009; Taylor & Brown, 1988, 1994). For example, people believe they are more virtuous, intelligent, talented, and compassionate than the average person (Alicke & Govorun, 2005; Batson & Collins, 2011). In fact, people exhibit a pronounced sense of moral superiority obtained by evaluating themselves more favorably than others (Batson & Collins, 2011; Tappin & McKay, 2017). And people use memory search, reconstruction, and interpretation to support more favorable evaluations of themselves than of others (D’Argembeau & Van der Linden, 2008; Demiray & Janssen, 2015; Kunda, 1990; Stanley & De Brigard, 2019; Stanley et al., 2017, 2019a). For example, people are more likely to recall and vividly reexperience positive information about themselves than they are to recall such information about others (D’Argembeau & Van der Linden, 2008). If people evaluate themselves more favorably than others, then they may excuse their own remembered transgressions more than similar kinds of remembered transgressions committed by others. In this way, moral principles would not always actively regulate moral judgments and decisions; instead, the accessibility of certain events in memory might dictate the endorsement of moral principles moment to moment.

In two studies, we investigate whether the accessibility of relevant events in memory shifts agreement with common moral principles. Using complementary experimental designs,

Studies 1 and 2 test the central hypothesis that, after recalling an event in which they personally violated a moral principle, participants would agree less strongly with the violated principle, relative to participants who recall others’ violations. Study 2 addresses the proposed mechanism with a mediation design. That is, Study 2 tests whether the shift in moral principle agreement after recalling a personal past violation of that principle is explained, in part, by participant’s willingness to selectively excuse their own immoral actions.

Both studies presented herein were formally preregistered. For both studies, we report all exclusion criteria, all materials and conditions included, and all independent and dependent measures. Deidentified data are publicly available on OSF (<https://osf.io/nqjf7/>).

## Study 1

In Study 1, participants recalled an event in which they violated a moral principle, or an event in which another person violated a moral principle. After recalling the event in accordance with the cue, they then rated their agreement with the violated moral principle. We assume that, because of random assignment to conditions, participants across conditions exhibit the same level of agreement, on average, with the moral principles prior to recalling a violation committed by themselves or by others. Assuming no baseline differences in moral principle agreement, any differences in agreement observed after recalling different kinds of past events is likely produced by change in principle agreement following the experimental manipulation. The preregistration for Study 1 is available on OSF (<https://osf.io/cq32m>). Study 2 then uses a pre–post experimental design to more directly measure change in moral principle agreement.

## Materials and methods

**Participants** A total of 1,205 individuals completed this study via Amazon’s Mechanical Turk for monetary compensation. Participant recruitment was restricted to individuals in the United States who had completed at least 50 HITs and had a prior approval rating of at least 90%. Ninety-three participants were excluded for failing a check question. As such, we were left with 1,112 participants ( $M_{\text{age}} = 36$  years,  $SD = 11$ , age range: 18–77, 502 females, 606 males). We conducted a power analysis using G\*Power for a one-way analysis of variance (ANOVA) with power ( $1 - \beta$ ) set at .85 and  $\alpha = .05$  (two-tailed). We aimed to recruit a total of 1,205 to detect small-to-moderate effect sizes for each of the three principles (after expected exclusions). This target sample size would ensure that we would have at least 100 participants per cell in our experimental design. All participants reported being fluent English speakers. Informed consent was obtained from each

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

**Materials** Three moral principles were in this study: (1) People should be honest with other people; (2) people should treat other people fairly; and (3) people should not harm other people. Recent research has successfully used cues involving dishonesty, unfairness, and harm to elicit memories of morally wrong actions committed by oneself and by others (Stanley et al., 2019a). Therefore, we expected that most participants would be able to recall a specific memory from the personal past in which they violated each of these principles.

**Procedure** This study was self-paced and consisted of a single session. Participants were randomly presented with one of the three moral principles and randomly assigned to the *actor condition*, the *recipient condition*, or the *observer condition*. In the actor condition, participants were asked to recall an event from their personal past in which they themselves violated the principle (i.e., dishonesty, unfairness, harm). In the recipient condition, participants were asked to recall an event from their personal past in which someone else violated the principle and the participant was the recipient of the transgression (e.g., a memory in which someone else was dishonest with the participant). In the observer condition, participants were asked to recall an event from their personal past in which they witnessed someone else violating the principle and the participant was not the recipient of the transgression. We included both recipient and observer conditions as comparisons to comprehensively account for the possible ways in which people actually encounter moral transgressions in the world. Moreover, the inclusion of both recipient and observer conditions also helps to rule out the possibility that the hypothesized self–other effects are simply the product of having been a direct participant in the remembered event (either as an actor or as a victim), as opposed to a mere witness.

Regardless of condition, participants were instructed to recall only actions that they believed to be morally wrong. All participants were also reminded that all of their responses were completely confidential. All participants described the remembered event in 3–5 sentences and provided the month and year that it occurred. This strategy for eliciting memories of moral transgressions was adapted from Stanley et al. (2019a). On the next page, participants were asked whether they were able to generate a memory of an action that they believe to be morally wrong in accordance with the cue provided. Participants who reported that they did not generate a memory in accordance with the cue were excluded from our analyses (a total of 19 participants reported that they did not generate a memory in accordance with the cue). Specifically, 98%, 98%, and 99% of participants reported being able to recall a morally wrong violation of the principle in the actor, recipient, and observer conditions, respectively.

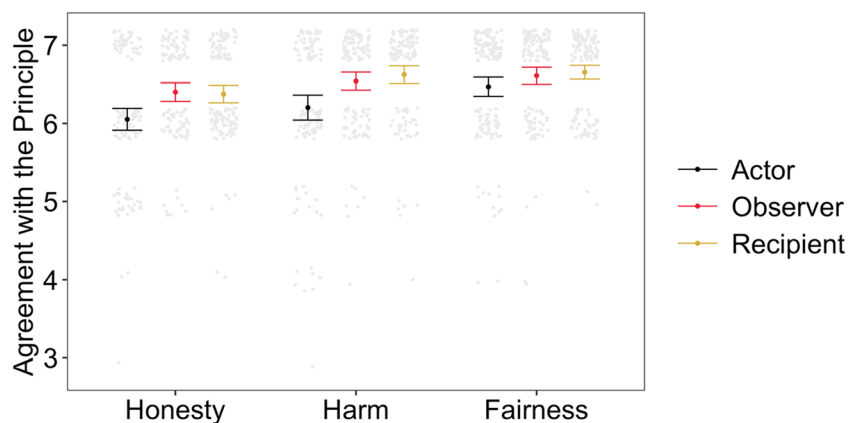
Participants then completed an unrelated 1-minute categorization distractor task, after which they were presented with the same moral principle again. Participants were instructed to make a judgment about how much they agreed or disagreed with the principle on a 7-point scale (1 = *strongly disagree*, 7 = *strongly agree*).

At the end, participants were asked the following: 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 did 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 (see exclusions, above). This same attention check question has been used in recent published research (e.g., Henne et al., 2019; Stanley et al., 2019b, 2020, 2021). Upon completion, participants were monetarily compensated for their time.

**Statistical analyses** Data were analyzed using R with the ‘lme4’ software package (Bates et al., 2015) and the ‘lmerTest’ software package (Kuznetsova et al., 2017). Data were fit to linear mixed-effects models (LMEM). 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 (on our view, 95% CIs around beta-values offer the best available indication of effect size for LMEMs). We present the results of LMEMs in the main text, but separate analyses for each individual principle are available in the supplement.

## Results

We tested our central hypothesis that, after recalling an event in which they personally violated a moral principle, participants would agree less strongly with the violated principle—relative to participants who recalled moral principle violations committed by others. To this end, we computed a LMEM with condition (actor, recipient, observer) on moral principle agreement ratings. The moral principle was the random effect in the model (random intercepts only). Corroborating our hypothesis, participants in the actor condition agreed less strongly with the moral principles than participants in the recipient ( $b = .31, SE = .05, t = 6.21, p < .001, 95\% CI [.21, .41]$ ) and observer conditions ( $b = .28, SE = .05, t = 5.57, p < .001, 95\% CI [.18, .38]$ ); however, there was no difference in the change in agreement ratings between recipient and observer conditions ( $p = .58$ ). Figure 1 graphically depicts agreement ratings for each individual moral principle as a function of experimental condition. See [Supplemental Information](#) for descriptive statistics and additional analyses for each principle taken separately.



**Fig. 1** Mean agreement with the moral principles. Error bars indicate 95% confidence intervals. Light-grey points represent individual participant responses evenly jittered

## Study 2

Study 2 builds on Study 1 by implementing a pre–post experimental design to directly measure *change* in moral principle agreement after recalling a moral principle violation committed by oneself or by others. Study 2 also tests whether the selective willingness for people to excuse their own immoral actions, but not the immoral actions of others, explains the shift in agreement with the moral principles. The preregistration for Study 2 is available on OSF (<https://osf.io/sb96tf>).

## Materials and methods

**Participants** A total of 1,510 individuals completed this study via Prolific for monetary compensation. Participant recruitment was restricted to individuals in the United States and the United Kingdom. One hundred and ninety-three participants were excluded for failing a check question (see below for details). As such, we were left with 1,317 participants ( $M_{\text{age}} = 35$  years,  $SD = 14$ , age range: 18–80, 794 females, 499 males). We determined the overall sample size based on the sample size (per cell) in Study 1. We aimed to recruit at least 1,500 to ensure that we would have at least 100 participants per cell in our experimental design (after expected exclusions). All participants reported being fluent English speakers. Informed consent was obtained from each participant in accordance with protocol approved by the Duke University Campus Institutional Review Board.

**Materials** Four moral principles were in this study: (1) People should be honest with other people; (2) people should not harm other people; (3) people should be loyal to other people; and (4) people should show respect to other people. Note that principles (1) and (2) were used in Study 1, but principles (3) and (4) are new to Study 2. We included these two new moral principles to obtain more variety in our set of moral principles,

as we expect our hypothesized effects to generalize across all moral principles.

**Procedure** This study was self-paced and consisted of a single session. Participants were randomly presented with one of the four moral principles, and they were asked to make an initial judgment about how much they agreed with the principle (1 = *strongly disagree*, 7 = *strongly agree*). Participants then completed an unrelated 1-min categorization distractor task, after which they were randomly assigned to the *actor condition*, the *recipient condition*, or the *observer condition*. The instructions for each condition were the same as in Study 1. As before, participants described the memory in 3–5 sentences and then reported the month and year that the event occurred. Using a 7-point scale (1 = *definitely no*, 7 = *definitely yes*), participants then indicated whether they believed that the circumstances surrounding the remembered event *excused* the moral principle violator’s action.

On the next page, participants were asked whether they were able to generate a memory of an action that they believe to be morally wrong in accordance with the cue provided. Participants who reported that they did not generate a memory in accordance with the cue were excluded from our analyses (a total of 171 participants reported that they did not generate a memory in accordance with the cue). Specifically, 82%, 91%, and 97% of participants reported being able to recall a morally wrong violation of the principle in the actor, recipient, and observer conditions, respectively.

Participants then completed an unrelated 1-minute categorization distractor task, after which they were presented with the same moral principle again. Participants were instructed to make a final judgment about how much they agreed or disagreed with the principle on the same 7-point scale (1 = *strongly disagree*, 7 = *strongly agree*). At the end, participants answered the same attention check question as in Study 1. Upon completion, participants were monetarily compensated for their time.

**Statistical analyses** The statistical software and methods used in Study 1 were also used in Study 2. To address possible mediation effects, using the ‘mediation’ package in R (Tingley et al., 2014), we also computed the average causal mediation effect (ACME), or the indirect effect, and the proportion of the effect mediated.

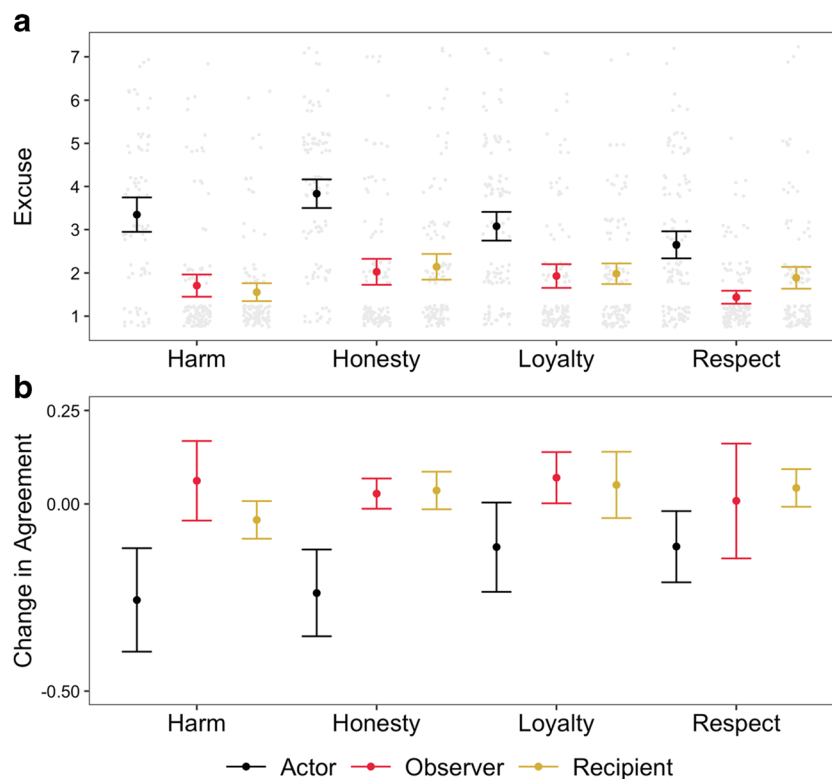
## Results

We tested our central hypothesis that, after recalling an event in which they personally violated a moral principle, participants would agree less strongly with the violated principle—relative to participants who recalled moral principle violations committed by others. To this end, we computed a LMEM with condition (actor, recipient, observer) on the change in agreement ratings (final rating *minus* initial rating). The moral principle was the random effect in the model (random intercepts only). Participants in the actor condition shifted principle agreement ratings (toward disagreeing from agreeing) more than participants in the recipient ( $b = .20, SE = .03, t = 5.72, p < .001, 95\% CI [.13, .26]$ ) and observer conditions ( $b = .22, SE = .03, t = 6.28, p < .001, 95\% CI [.14, .29]$ ); however, there was no difference in the change in agreement ratings between recipient and observer conditions ( $p = .55$ ). Figure 2 graphically depicts change in agreement ratings for each individual moral principle as a function of experimental condition. See

[Supplemental Information](#) for descriptive statistics and analyses with each principle taken separately.

Next, we investigated whether participants are more likely to excuse their own moral principle violations, relative to the moral principle violations of others. To this end, we computed a LMEM with condition (actor, recipient, observer) on excused judgments. The moral principle was the random effect in the model. Participants in the actor condition were more likely to report that their moral principle violations were excused than participants in the recipient ( $b = 1.31, SE = .10, t = 12.95, p < .001, 95\% CI [1.12, 1.50]$ ) and observer conditions ( $b = 1.43, SE = .10, t = 14.08, p < .001, 95\% CI [1.24, 1.63]$ ); however, there was no difference in excused judgments between the recipient and observer conditions ( $p = .22$ ). Figure 2 graphically depicts excused judgments for each individual principle as a function of experimental condition.

Finally, we tested whether the effect of condition on the change in moral principle agreement judgments is mediated by beliefs about the transgressor being excused. To this end, we computed two mediation models within a LMEM framework. In the first model, only actor and recipient conditions were included. Beliefs about the transgressor being excused mediated the effect of condition (actor vs. recipient) on the change in agreement judgments (ACME = .07,  $p < .001, 95\% CI [.03, .10]$ ; Prop. Mediated = .31,  $p < .001, 95\% CI [.16, .58]$ ). In the second model, only actor and observer conditions



**Fig. 2** Mean excused judgments (a) and change in moral principle agreement (final judgment *minus* initial judgment) (b) for each moral principle as a function of condition. Error bars indicate 95% confidence intervals. Light-grey points represent individual participant responses evenly jittered



were included. Beliefs about the transgressor being excused mediated the effect of condition (actor vs. observer) on the change in agreement ratings (ACME = .07,  $p < .001$ , 95% CI [.04, .10]; Prop. Mediated = .34,  $p < .001$ , 95% CI [.20, .56]).

## General discussion

Across two studies, we investigated whether recalling certain past events situationally shifts agreement with commonly endorsed moral principles. Both studies corroborate our central hypothesis: relative to participants who recalled an instance in which another person violated a moral principle, when participants recalled an instance in which they personally violated a moral principle, they agreed less strongly, on average, with the very principle they violated. Interestingly, most participants continued to agree with the moral principles after recalling their own violations, but they agreed less strongly than participants who recalled violations committed by others, regardless of whether the participants themselves were the victims of such violations. In other words, participants did not tend to reject the moral principles after recalling their own violations, but they did hold those principles with less conviction than other participants who recalled other people's moral violations. In Study 2, we then found that participants were more willing to excuse their own remembered violations than others' violations. This selective willingness for participants to excuse their own violations, but not the violations of others (whether or not they were the victims), statistically accounted for the observed shifts in moral principle agreement.

Our findings are compatible with the literature on moral disengagement, but they do diverge from this past research in important ways. Moral disengagement mechanisms (e.g., distorting the consequences of actions, dehumanizing victims) help people to convince themselves that their actions are permissible and that their ethical standards need not apply in certain contexts (Bandura, 1999; Bandura et al., 1996; Detert et al., 2008). These disengagement mechanisms are thought to help people to protect their favorable views of themselves. Note that convincing oneself that a particular action is morally acceptable in a particular context via moral disengagement entails maintaining the same level of agreement with the overarching moral principles; the principle just does not apply in some particular context. In contrast, our findings suggest that by reflecting on their own morally objectionable actions, people's agreement with the overarching, guiding principles changes. It is not that the principle does not apply; it is that the principle is held with less conviction.

Empirical research on human morality has largely focused on how and why individuals make certain moral judgments and decisions using hypothetical vignettes (e.g., Cushman et al., 2006; Greene et al., 2001; Nichols & Mallon, 2006;

Henne et al., 2016; Henne & Sinnott-Armstrong, 2018; Stanley et al., 2018; Valdesolo & DeSteno, 2006). These studies have produced valuable insights into how people make moral judgments and decisions. However, people rarely, if ever, experience events that are identical, or even similar, to the contrived dilemmas that have dominated the moral psychology literature (FeldmanHall et al., 2012; Hofmann et al., 2014). In fact, recent research has suggested that judgments and decisions made in hypothetical dilemmas do not predict how people make real-life moral judgments and decisions (Bostyn et al., 2018). A key advantage of the current research is the use of ecologically valid and personally relevant memories of real-life transgressions committed by the participants and by others.

Normative ethical theories and religious traditions that offer general moral principles are meant to help us to understand aspects of ourselves and our world in ways that offer insights and guidance for living a moral life (Albertzart, 2013; Väyrynen, 2008). Our findings introduce some cause for doubt about the stability of moral principles over time, and therefore, their reliability as accurate indicators of moral judgments and actions in the real world. Future research should investigate whether and to what extent the effects identified in our studies generalize to other kinds of moral violations. For example, while our effects are clear and consistent for harm, honesty, and loyalty violations, people may be less willing to excuse violations that elicit strong affective responses like disgust (violations of purity and sanctity norms)—regardless of who commits the transgression.

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.3758/s13423-021-01935-8>.

## Declarations

**Conflict of interest** The authors declared that they have no competing interests with respect to the publication of this article.

## References

- Albertzart, M. (2013). Principle-based moral judgment. *Ethical Theory and Moral Practice*, 16, 339–354. <https://doi.org/10.1007/s10677-012-9343-x>
- Alicke M. D., & Govorun, O. (2005). The better-than-average effect. In M. D. Alicke, D. A. Dunning, & J. I. Krueger (Eds.), *The self in social judgment* (pp. 85–106). 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, 1–48. <https://doi.org/10.1080/10463280802613866>
- Bandura, A. (1999). Moral disengagement in the perpetration of inhumanities. *Personality and Social Psychology Review*, 3(3), 193–209. [https://doi.org/10.1207/s15327957pspr0303\\_3](https://doi.org/10.1207/s15327957pspr0303_3)

- Bandura, A., Barbaranelli, C., Caprara, G. V., & Pastorelli, C. (1996). Mechanisms of moral disengagement in the exercise of moral agency. *Journal of Personality and Social Psychology, 71*(2), 364–375. <https://doi.org/10.1037/0022-3514.71.2.364>
- Bates, D., Maechler, M., & Bolker, B. (2015). Walker. S. lme4: Linear mixed-effects models using S4 classes. *Journal of Statistical Software, 67*, 1–48.
- Batson, C. D., & Collins, E. C. (2011). Moral hypocrisy: A self-enhancement/self-protection motive in the moral domain. In M. D. Alicke & C. Sedikides (Eds.), *Handbook of self-enhancement and self-protection* (pp. 92–111). Guilford Press.
- 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*, 1084–1093. <https://doi.org/10.1177/0956797617752640>
- Cushman, F., & Young, L. (2011). Patterns of moral judgment derive from nonmoral psychological representations. *Cognitive Science, 35*(6), 1052–1075. <https://doi.org/10.1111/j.1551-6709.2010.01167.x>
- Cushman, F., Young, L., & Hauser, M. (2006). The role of conscious reasoning and intuition in moral judgment: Testing three principles of harm. *Psychological Science, 17*(12), 1082–1089. <https://doi.org/10.1111/j.1467-9280.2006.01834.x>
- D'argembeau, A., & Van der Linden, M. (2008). Remembering pride and shame: Self-enhancement and the phenomenology of autobiographical memory. *Memory, 16*(5), 538–547.
- Demiray, B., & Janssen, S. M. (2015). The self-enhancement function of autobiographical memory. *Applied Cognitive Psychology, 29*(1), 49–60.
- Detert, J. R., Treviño, L. K., & Sweitzer, V. L. (2008). Moral disengagement in ethical decision making: A study of antecedents and outcomes. *Journal of Applied Psychology, 93*(2), 374–391. <https://doi.org/10.1037/0021-9010.93.2.374>
- 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*, 434–441. <https://doi.org/10.1016/j.cognition.2012.02.001>
- Greene, J. D., Sommerville, R. B., Nystrom, L. E., Darley, J. M., & Cohen, J. D. (2001). An fMRI investigation of emotional engagement in moral judgment. *Science, 293*(5537), 2105–2108. <https://doi.org/10.1126/science.1062872>
- Henne, P., Chituc, V., De Brigard, F., & Sinnott-Armstrong, W. (2016). An empirical refutation of 'ought' implies 'can'. *Analysis, 76*(3), 283–290.
- Henne, P., Niemi, L., Pinillos, Á., De Brigard, F., & Knobe, J. (2019). A counterfactual explanation for the action effect in causal judgment. *Cognition, 190*, 157–164. <https://doi.org/10.1016/j.cognition.2019.05.006>
- Henne, P., & Sinnott-Armstrong, W. (2018). Does neuroscience undermine morality? In G. D. Caruso & O. Flanagan (Eds.), *Neuroexistentialism: Meaning, morals, and purpose in the age of neuroscience*. Oxford University Press.
- Hofmann, W., Wisneski, D. C., Brandt, M. J., & Skitka, L. J. (2014). Morality in everyday life. *Science, 345*(6202), 1340–1343. <https://doi.org/10.1126/science.1251560>
- Horne, Z., Powell, D., & Hummel, J. (2015). A single counterexample leads to moral belief revision. *Cognitive Science, 39*(8), 1950–1964. <https://doi.org/10.1111/cogs.12223>
- Kant, I. (1959). *Foundation of the metaphysics of morals* (L. W. Beck, Trans.). Bobbs-Merrill. (Original work published 1785)
- Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin, 108*(3), 480–498. <https://doi.org/10.1037/0033-2909.108.3.480>
- Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. (2017). lmerTest package: tests in linear mixed effects models. *Journal of Statistical Software, 82*(13), 1–26.
- Lombrozo, T. (2009). The role of moral commitments in moral judgment. *Cognitive Science, 33*(2), 273–286. <https://doi.org/10.1111/j.1551-6709.2009.01013.x>
- Mill, J. S. (1998). *Utilitarianism* (R. Crisp, Ed.) Oxford University Press. (Original work published 1861)
- Nichols, S., & Mallon, R. (2006). Moral dilemmas and moral rules. *Cognition, 100*, 530–542. <https://doi.org/10.1016/j.cognition.2005.07.005>
- Stanley, M. L., & De Brigard, F. (2019). Moral memories and the belief in the good self. *Current Directions in Psychological Science, 28*(4), 387–391. <https://doi.org/10.1177/0963721419847990>
- Stanley, M. L., Dougherty, A. M., Yang, B. W., Henne, P., & De Brigard, F. (2018). Reasons probably won't change your mind: The role of reasons in revising moral decisions. *Journal of Experimental Psychology: General, 147*, 962–987. <https://doi.org/10.1037/xge0000368>
- Stanley, M. L., Henne, P., & De Brigard, F. (2019a). Remembering moral and immoral actions in constructing the self. *Memory & Cognition, 47*(3), 441–454. <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. <https://doi.org/10.1037/xge0000317>
- Stanley, M. L., Marsh, E. J., & Kay, A. C. (2020). Structure-seeking as a psychological antecedent of beliefs about morality. *Journal of Experimental Psychology: General, 149*, 1908–1918. <https://doi.org/10.1037/xge0000752>
- Stanley, M. L., Stone, A. R., & Marsh, E. J. (2021). Cheaters claim they knew the answers all along. *Psychonomic Bulletin & Review, 28*, 341–350. <https://doi.org/10.3758/s13423-020-01812-w>
- Stanley, M. L., Yin, S., & Sinnott-Armstrong, W. (2019b). A reason-based explanation for moral dumbfounding. *Judgment and Decision Making, 14*(2), 120–129.
- Tappin, B. M., & McKay, R. T. (2017). The illusion of moral superiority. *Social Psychological and Personality Science, 8*(6), 623–631. <https://doi.org/10.1177/1948550616673878>
- Taylor, S. E., & Brown, J. D. (1988). Illusion and well-being: A social psychological perspective on mental health. *Psychological Bulletin, 103*, 193–210. <https://doi.org/10.1037/0033-2909.103.2.193>
- Taylor, S. E., & Brown, J. D. (1994). Positive illusions and well-being revisited: Separating fact from fiction. *Psychological Bulletin, 116*, 21–27. <https://doi.org/10.1037/0033-2909.116.1.21>
- Tingley, D., Yamamoto, T., Hirose, K., Keele, L., & Imai, K. (2014). Mediation: R package for causal mediation analysis. *Journal of Statistical Software, 59*, 5. <https://doi.org/10.18637/jss.v059.i05>
- Valdesolo, P., & DeSteno, D. (2006). Manipulations of emotional context shape moral judgment. *Psychological Science, 17*(6), 476–477. <https://doi.org/10.1111/j.1467-9280.2006.01731.x>
- Väyrynen, P. (2008). Usable moral principles. In V. Strahovnik, M. Potrc, & M. N. Lance (Eds.), *Challenging moral particularism* (pp. 75–102). Routledge.