

Reflecting on Identity Change Facilitates Confession of Past Misdeeds

Beth Anne Helgason¹

Jonathan Zev Berman²

¹ Organisational Behaviour, London Business School

² Marketing, London Business School

Article in Press at *Journal of Experimental Psychology: General*

January 31st, 2022

Corresponding author: Beth Anne Helgason (bhelgason@london.edu)

Address: London Business School, Sussex Place, Regent's Park, London NW1 4SA, England

Data files for all studies can be accessed at: <https://osf.io/yax8z/>

Abstract

Across four studies ($N = 3,351$), we demonstrate that reflecting on identity change increases confession and decreases justification of past misdeeds. Moreover, publicly communicating one's identity change to others increases confession above and beyond privately reflecting on identity change. By severing their connection with their past self, individuals can admit to past a misdeed ("I did it"), while reducing their fear that doing so will implicate their present moral character ("But that's not who I am anymore").

Keywords: *Confession, Identity change, Identity continuity, Unethical behavior, Moral disengagement*

Confessing to a misdeed not only helps right a wrong, but also provides a deep sense of relief (Kassin & Gudjonsson, 2004; Peer et al., 2014; Pennebaker, 1997). Yet, people are reluctant to confess—even when the risk of punishment is negligible (Peer et al., 2014). If “the truth shall set you free,” then why is it so hard for people to admit to having sinned in their past? And what might enable people to own up to their past misdeeds?

We argue that one reason why people fail to confess is because they fear that admitting to past misdeeds will taint their present moral character. People have a strong desire to maintain a reputation as a good and moral person both to themselves (Aquino & Reed, 2002; Monin & Jordan, 2009) and to others (Banerjee et al., 2020; Goffman, 1959; Vonasch et al., 2018). Confessing to a past misdeed can threaten that reputation. Transgressions are viewed as particularly relevant to moral character, much more so than virtuous behavior (Birnbaum, 1972; Skowronski & Carlston, 1989). Moreover, because moral character is considered to be a particularly consistent and enduring aspect of personal identity (Strohming et al., 2017; Strohming & Nichols, 2014), even wrongdoings from the distant past can leave a lasting stain on an individual’s reputation (Berman & Small, 2018; Molouki & Bartels, 2017).

In this research, we examine whether prompting individuals to reflect on how they have changed over time will facilitate confession and reduce justification of their past misdeeds. We propose that reflecting on identity change enables people to sever the connection between their past self (the one that acted immorally) and their present self (the one they would like to see as moral). Rather than viewing their past actions as a stain on their present character, reflecting on identity change may allow people to construe their former transgression as a signal of personal growth (Schumann & Dweck, 2014; Stanley et al., 2017; Wilson & Ross, 2003).

Our predictions are consistent with a growing body of research suggesting that identity-discontinuity mitigates the threat that unethical behaviors pose to one's sense of self. For instance, people are more likely to engage in unethical acts in the present the more they experience identity-discontinuity with their future self (Hershfield et al., 2012). Moreover, people are motivated to create a sense of discontinuity with their past unethical actions (Ross & Wilson, 2003; Stanley et al., 2019; Wilson & Ross, 2003). They recall their moral failures as occurring further back in time than their moral triumphs (Escobedo & Adolphs, 2010), and recall engaging in worse moral transgressions during times in which they felt different from who they are now (Stanley et al., 2017). Yet it remains unclear whether an intervention prompting people to reflect on past identity change can increase confession and reduce justification of past misdeeds.

In sum, we investigate whether reflecting on identity change increases confession (Studies 1-3) and reduces justification (Study 4) of past misdeeds, and we examine the role of private self-image and public reputation-management concerns in this effect (Study 3). Studies 1-3 were pre-registered (<https://osf.io/yax8z/>). Any measures not reported in the main manuscript are presented in the Online Supplement. Study procedures were approved under research ethics protocol #REC726-24032024 at the authors' institution.

Study 1

In Study 1, we examine whether participants who reflected on how they have fundamentally changed (vs. stayed the same) over time would be more likely to voluntarily confess to a past misdeed.

Method

Participants

We aimed to recruit 800 participants from MTurk and collected a total of 816 participants. Per our pre-registration, we removed 64 participants who failed an attention check, resulting in a final sample of 752 participants (428 women, 318 men, 6 other, $M_{\text{age}} = 38$).

Procedure

We randomly assigned participants to either write in a text box how they have changed (Identity-change condition) or stayed the same (Identity-continuity condition) since high school in a between-subjects design. Participants were then asked if they would like to confess to a past-misdeed (our main DV). Those who selected “yes” were provided a free-text box to write their confession.

Next, participants reported how much they have fundamentally changed since high school; and to what extent the person they are now is similar to the person they were in high school (reverse-coded), each on seven-point scales. These were averaged to create a two-item measure of perceived identity change ($r = .66$, $p < .001$). Finally, participants responded to an attention check (see Online Supplement) and reported their demographics.

Results

Confession Rate

Supporting our hypothesis, a greater proportion of participants confessed to a past misdeed after reflecting on how they had changed over time (36.6%) than after reflecting on how they had remained the same (27.5%), $\chi^2(1, N = 752) = 7.11$, $p = .008$, $\phi = .10$.

Perceived Identity Change

Participants in the Identity-change condition reported having changed more since high school ($M = 5.39$, $SD = 1.30$) than those in the Identity-continuity condition ($M = 4.14$, $SD = 1.45$), $t(750) = 12.46$, $p < .001$, 95% CI_d [0.76, 1.06]. Furthermore, perceived identity change mediated the effect of condition on confession, $b = 3.49$, 95% CI [2.87, 4.25], when computing the indirect effect in a bootstrapped mediation analysis.

Study 2

Study 1 found that reflecting on one's past identity change increases confession relative to reflecting on identity continuity. However, it is unclear whether reflecting on identity change facilitates confession, reflecting on identity continuity reduces confession, or both. To evaluate the directionality of the effect we included a control condition in Study 2.

Method

Participants

We aimed to recruit 1,000 participants from MTurk and collected a total of 1,012 participants. Per our pre-registration, we removed 73 participants who failed an attention check, resulting in a final sample of 939 participants (547 women; 388 men; 4 did not disclose; $M_{\text{age}} = 37$).

Procedure

We randomly assigned participants to either to an Identity-change, an Identity-continuity, or a Control condition in a between-subjects design. Participants in the Identity-change and Identity-continuity conditions responded to the same prompts as in Study 1. Participants in the Control condition wrote about what they did the past weekend.

Next, participants had the opportunity to confess to six transgressions commonly committed during high school (e.g., copying homework, disrespecting a teacher). We operationalized confession as the degree to which participants admitted to engaging in the misdeeds from 1 = “Never” to 5 = “Often”. Following our pre-registration, we averaged these items to create a six-item measure of confessed misdeeds ($\alpha = .76$).

Finally, participants rated their perceived identity change on the same measure as in Study 1 ($r = .58$, $p < .001$), responded to an attention check (see Online Supplement), and reported their demographics.

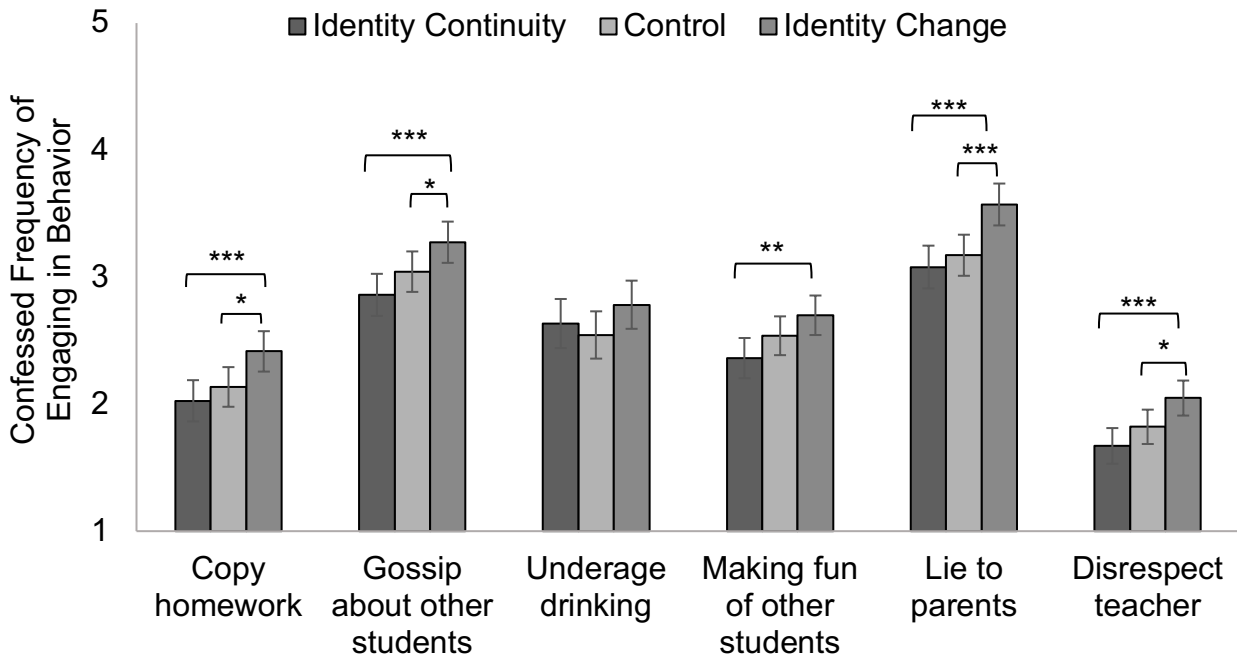
Results

Confession of Past Transgressions

Participants in the Identity-change condition confessed to transgressing more frequently ($M = 2.19$, $SD = 0.62$) than those in the Identity-continuity condition ($M = 2.01$, $SD = 0.53$), $t(612) = 4.02$, $p < .001$, 95% CI_d [0.16, 0.48] and those in the Control condition ($M = 2.07$, $SD = 0.54$), $t(636) = 2.86$, $p = .004$, 95% CI_d [0.04, 0.22]. The difference between the Identity-continuity and Control conditions was not significant, $t(624) = -1.22$, $p = .223$, 95% CI_d [-0.26, 0.05]. Figure 1 displays the results by individual transgression.

Figure 1

Confessed Frequency of Engaging in each Behavior in Study 2



Note. The confessed frequency of engaging in each behavior ranges from 1 = “Never” to 5 = “Often”. Error bars correspond to 95% confidence intervals.

*** $p < .001$, ** $p < .01$, * $p < .05$

Perceived Identity Change

Participants in the Identity-change condition reported having changed more since high school ($M = 5.44$, $SD = 1.29$) than those in the Identity-continuity condition ($M = 4.28$, $SD = 1.35$), $t(612) = 10.43$, $p < .001$, 95% CI_d [0.71, 1.04], and those in the Control condition ($M = 5.09$, $SD = 1.45$), $t(636) = 3.18$, $p = .002$, 95% CI_d [0.09, 0.41]. Moreover, those in the Control condition reported having changed more since high school than those in the Identity-continuity condition, $t(624) = 7.38$, $p < .001$, 95% CI_d [0.42, 0.74].

We conducted a bootstrapped mediation analysis with condition as the categorical independent variable (1 = Identity-continuity, 2 = Control, 3 = Identity-change), perceived identity change as the mediator, and confession as the dependent variable (Hayes & Preacher, 2014). Perceived identity change significantly mediated the effect of condition on confession, $b = 0.21$, 95% CI [0.14, 0.27], for the total indirect effect, obtained by summing the indirect effects for the two contrasts (Control vs. Identity-Continuity; Identity-Change vs. Identity-Continuity).

Study 3

Study 3 extends our investigation in two ways. First, we examined whether reflecting on identity change facilitates confessing to a person whom participants might interact with in the future. Second, we examined the role of *communicating* identity change to others in facilitating confession. If people believe that public knowledge of their identity change will protect them from negative character judgments associated with past transgressions, then communicating identity change to others may be particularly effective in encouraging confession. Study 3 examined this possibility.

Method

Participants

We aimed to recruit 1,200 participants from Prolific Academic and collected a total of 1,212 participants. Per our pre-registration, we removed 31 participants who failed an attention check, resulting in a final sample of 1,182 participants (943 women; 225 men; 4 third gender; $M_{\text{age}} = 28$).

Procedure

We randomly assigned participants to either to a Private-reflection, Public-communication, or Control condition in a between-subjects design. Participants in the Private-reflection condition first wrote privately about how they have fundamentally changed since high school and were then paired with a partner (“Player B”). Participants in the Public-communication condition were first paired with a partner and then wrote *to their partner* about how they have changed since high school. Finally, participants in the Control Condition were paired with a partner and then wrote to their partner about what they did last weekend.

Next, participants in all conditions learned that they had the opportunity to confess to their partner about transgressions commonly committed during high school. Their partner would judge whether they would like to work with them in a future study based on their responses. We used a similar six-item measure of confessed misdeeds as Study 2 ($\alpha = .75$, see Figure 2 for items).

Thus, all participants confessed past misdeeds to their partner, but the three conditions differed in whether participants privately reflected on identity change, communicated their identity change to their partner, or did not consider identity change at all. Participants also responded to the same measure of perceived identity change as in Study 1 ($r = .59$, $p < .001$) and a single-item manipulation check of public versus private expression of identity change (“How much does Player B know about what you were like in high school?”).

Results

Public versus Private Expression of Identity Change

Consistent with the intent of the manipulation, participants in the Public-communication condition rated their partner as knowing more about what they were like in high school ($M = 3.11$, $SD = 1.65$) than those in the Private-reflection condition ($M = 2.79$, $SD = 1.72$), $t(769) = 2.73$, $p = .006$, 95% CI_d [0.05, 0.33], and those in the Control condition ($M = 2.21$, $SD = 1.50$), $t(785) = 7.80$, $p < .001$, 95% CI_d [0.43, 0.72].

Confession of Past Transgressions

Participants in the Public-communication condition confessed to transgressing more frequently in high school ($M = 2.31$, $SD = 0.69$) than both those in the Control condition ($M = 2.06$, $SD = 0.56$), $t(785) = 5.64$, $p < .001$, 95% CI_d [0.26, 0.54] and those in the Private-reflection condition ($M = 2.19$, $SD = 0.61$), $t(769) = 2.51$, $p = .012$, 95% CI_d [0.03, 0.31]. Participants in the Private-reflection condition also confessed to transgressing more frequently than those in the Control condition, $t(806) = 3.14$, $p = .002$, 95% CI_d [0.19, 0.37]. Figure 2 displays the results by individual transgression.

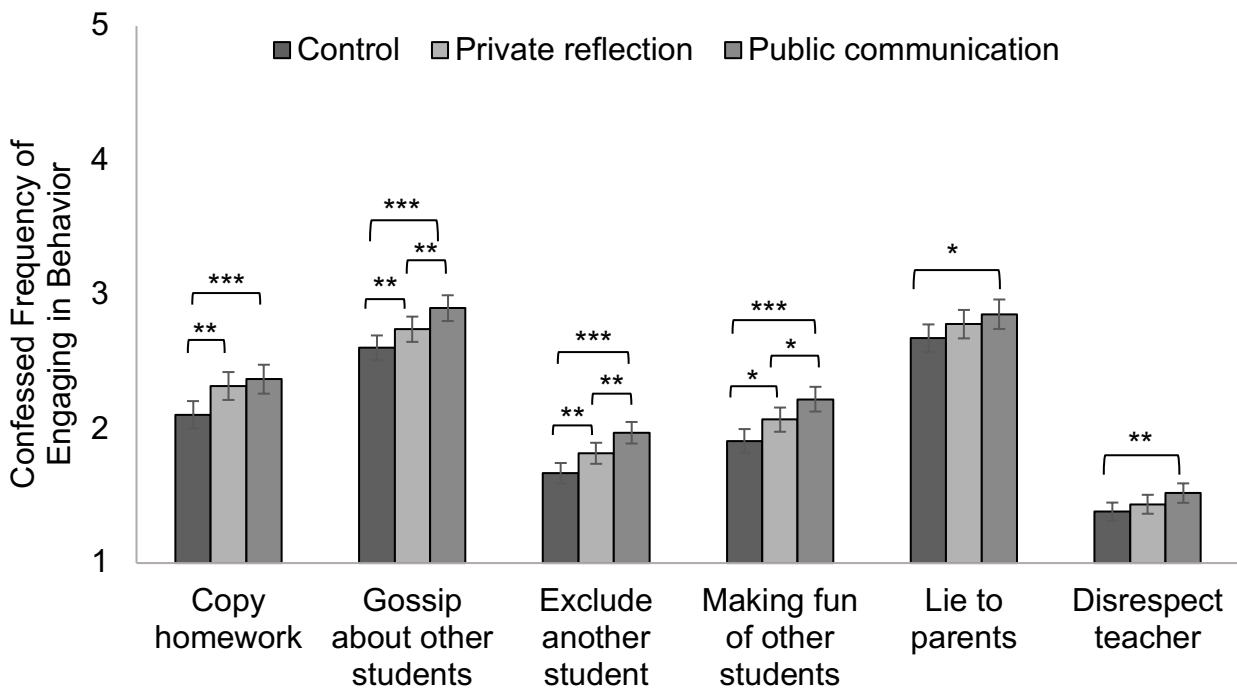
Perceived Identity Change

Contrary to our previous studies, we did not find evidence that participants in the Public-communication condition reported having changed more since high school ($M = 5.54$, $SD = 1.13$) than participants in the Control condition ($M = 5.47$, $SD = 1.26$), $t(785) = 0.78$, $p = .437$, 95% CI_d [-0.08, 0.20], or that participants in the Private-reflection condition reported having changed

more ($M = 5.51$, $SD = 1.21$) than those in the Control condition, $t(769) = 0.48$, $p = .631$, 95% CI_d [-0.11, 0.17].¹

Figure 2

Confessed Frequency of Engaging in Each Behavior in Study 3



Note. The confessed frequency of engaging in each behavior ranges from 1 = “Never” to 5 = “Often”. Error bars correspond to 95% confidence intervals.

*** $p < .001$, ** $p < .01$, * $p < .05$

Study 4

¹ Given that we find significant effects on perceived identity change in all other studies (including two supplemental studies), random error variance may account for this null effect. A second possibility is that an additional mechanism mediates the effect of the identity change manipulation.

Studies 1-3 show that reflecting on identity change increases confession. Past research reveals that even when people do confess, they often excuse or justify their misbehaviors (Bandura et al., 1996; Shalvi et al., 2015). In Study 4, we examine whether reflecting on identity change also reduces people's tendencies to justify their past misdeeds.

Method

Participants

We aimed to recruit 600 participants from MTurk and collected a total of 613 participants for this study. We removed responses from 135 participants who failed at least one of our attention checks, leaving a final sample of 478 participants (201 men, 257 women, 20 did not disclose; $M_{\text{age}} = 35$).

Procedure

Participants first indicated one misdeed that they had a specific memory of doing in high school from the following options: lied to a friend or family member (32.0%); gossiped, spread rumours about, or teased another student (13.2%); stole something (12.9%); started a fight (8.7%); mistreated a teacher or parent (7.6%); or bullied or excluded another student (6.5%). We gave participants the option to indicate if they had no memories of engaging in any of these behaviors, in which case they finished the study after this question (19.2%).

Participants who indicated a specific misdeed were randomly assigned to either reflect on their identity change or identity continuity with the same prompts as in Studies 1-3, in a two-cell between-subjects design. Following their reflection, participants described their past misdeed in a few sentences in a free-response text box.

As our main outcome measure, participants responded to a four-item measure of moral disengagement (adapted from Bandura et al., 1996). Specifically, participants reported their agreement with the following statements: “I shouldn't be blamed for what I did because many others do it too”; “What I did is no big deal when you consider what others do in high school”; “It is okay to do what I did because other things that people do are much worse”; and “What I did was okay because it didn't really do any harm.” (from -3 = “Strongly disagree” to 3 = “Strongly agree;” $\alpha = .91$). Finally, participants responded to the same measure of identity change as in Studies 1-3 ($r = .65, p < .001$) and reported their demographics.

Results

Moral Disengagement

Supporting our predictions, participants in the Identity-change condition justified their past misdeeds less ($M = -1.05, SD = 1.64$) than participants in the Identity-continuity condition ($M = -0.68, SD = 1.55$), $t(462) = -2.49, p = .013, 95\% CI_d [-0.42, -0.05]$.

Perceived Identity Change

Participants in the Identity-change condition reported having changed significantly more since high school ($M = 5.77, SD = 1.19$) than those in the Identity-continuity condition ($M = 4.43, SD = 1.47$), $t(462) = 10.81, p < .001, 95\% CI_d [0.80, 1.19]$. Further, perceived identity change significantly mediated the effect of condition on moral disengagement, $b = -0.46, 95\% CI [-0.62, -0.29]$, in a bootstrapped mediation analysis.

General Discussion

Four studies show that prompting individuals to reflect on how they have changed over time increases their willingness to confess and reduces their tendencies to justify past misdeeds. We additionally find that publicly communicating identity change to others increases confession above and beyond simply reflecting on identity change. Together, these findings illustrate how personal narratives of change can help individuals come to terms with their past misdeeds, and advances research on the social and personal benefits of identity-discontinuity (Kim & Wohl, 2015; Schanbacher et al., 2021; Wohl et al., 2018).

Our findings have several limitations. For one, our studies were conducted with American samples that may not generalize to other cultural contexts (cf. Henrich et al., 2010). Additionally, most participants confessed to relatively minor misdeeds. Future research should examine whether our findings extend to confessing severe transgressions, particularly if the risk of punishment increases with transgression severity. Moreover, future research can examine whether our findings extend to confessing non-moral failings (e.g., failing a test).

In Study 3, we found that individuals who publicly communicated identity change were more willing to confess than those who privately reflected on identity change. This suggests that people believe that communicating their identity change will reduce the reputational costs associated with transgressions. However, it remains unclear whether these beliefs are accurate. Because identity change is hard to verify, observers may be suspicious of individuals who evoke identity change alongside confessing. Nonetheless, it is likely that a transgressor who professes identity change while *voluntarily* confessing will be believed more than a transgressor who professes identity change upon being exposed (cf. Weiner et al., 1991). Additionally, providing

costly signals of identity change (e.g., suffering for past wrongdoings; Effron & Miller, 2015) may increase the credibility of their claims.

In conclusion, we find that individuals are more willing to confess to their past unethical behaviors when they acknowledge how they have changed as a person over time. Doing so may allow individuals to confess past misdeeds (“I did it”), while reducing the fear that their past misdeeds will implicate their present moral character (“But that’s not who I am anymore”).

References

- Aquino, K., & Reed, A., II. (2002). The self-importance of moral identity. *Journal of Personality and Social Psychology*, 83(6), 1423–1440. <https://doi.org/10.1037//0022-3514.83.6.1423>
- 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), 364–374. <https://doi.org/10.1037/0022-3514.71.2.364>
- Banerjee, R., Heyman, G. D., & Lee, K. (2020). The development of moral self-presentation. In L. A. Jensen (Ed.), *The Oxford handbook of moral development: An interdisciplinary perspective* (pp. 92–109). Oxford University Press.
- Berman, J. Z., & Small, D. A. (2018). Discipline and desire: On the relative importance of willpower and purity in signaling virtue. *Journal of Experimental Social Psychology*, 76, 220–230. <https://doi.org/10.1016/j.jesp.2018.02.007>
- Birnbaum, M. H. (1972). Morality judgments: Tests of an averaging model. *Journal of Experimental Psychology*, 93(1), 35–42. <https://doi.org/10.1037/h0032589>
- Effron, D. A., & Miller, D. T. (2015). Do as I say, not as I've done: Suffering for a misdeed reduces the hypocrisy of advising others against it. *Organizational Behavior and Human Decision Processes*, 131, 16–32. <https://doi.org/10.1016/j.obhdp.2015.07.004>
- Escobedo, J. R., & Adolphs, R. (2010). Becoming a better person: Temporal remoteness biases autobiographical memories for moral events. *Emotion*, 10(4), 511–518. <https://doi.org/10.1037/a0018723>
- Goffman, E. (1959). *The Presentation of Self in Everyday Life*. Doubleday.

- Hayes, A. F., & Preacher, K. J. (2014). Statistical mediation analysis with a multicategorical independent variable. *British Journal of Mathematical and Statistical Psychology*, 67(3), 451–470. <https://doi.org/10.1111/bmsp.12028>
- Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences*, 33(2–3), 61–83. <https://doi.org/10.1017/S0140525X0999152X>
- Hershfield, H. E., Cohen, T. R., & Thompson, L. (2012). Short horizons and tempting situations: Lack of continuity to our future selves leads to unethical decision making and behavior. *Organizational Behavior and Human Decision Processes*, 117(2), 298–310. <https://doi.org/10.1016/j.obhdp.2011.11.002>
- Kassin, S. M., & Gudjonsson, G. H. (2004). The psychology of confessions: A review of the literature and issues. *Psychological Science in the Public Interest*, 5(2), 33–67. <https://doi.org/10.1111/j.1529-1006.2004.00016.x>
- Kim, H. S. (Andrew), & Wohl, M. J. A. (2015). The bright side of self-discontinuity: Feeling disconnected with the past self increases readiness to change addictive behaviors (via nostalgia). *Social Psychological and Personality Science*, 6(2), 229–237. <https://doi.org/10.1177/1948550614549482>
- Molouki, S., & Bartels, D. M. (2017). Personal change and the continuity of the self. *Cognitive Psychology*, 93, 1–17. <https://doi.org/10.1016/j.cogpsych.2016.11.006>
- Monin, B., & Jordan, A. H. (2009). The dynamic moral self: A social psychological perspective. In D. Narvaez & D. K. Lapsley (Eds.), *Personality, identity, and character* (pp. 341–354). Cambridge University Press. <https://doi.org/10.1017/CBO9780511627125.016>

- Peer, E., Acquisti, A., & Shalvi, S. (2014). “I cheated, but only a little”: Partial confessions to unethical behavior. *Journal of Personality and Social Psychology*, 106(2), 202–217.
<https://doi.org/10.1037/a0035392>
- Pennebaker, J. W. (1997). *Opening up: The healing power of expressing emotions*. Guilford Press.
- Ross, M., & Wilson, A. E. (2003). Autobiographical memory and conceptions of self: Getting better all the time. *Current Directions in Psychological Science*, 12(2), 66–69.
<https://doi.org/10.1111/1467-8721.01228>
- Schanbacher, A. D., Gurdamar-Okutur, N., & Faro, D. (2021). It’s no longer “me”: Low past-self-continuity reduces the sunk-cost bias. *Journal of Experimental Social Psychology*, 95, 104–146. <https://doi.org/10.1016/j.jesp.2021.104146>
- Schumann, K., & Dweck, C. S. (2014). Who accepts responsibility for their transgressions? *Personality and Social Psychology Bulletin*, 40(12), 1598–1610.
<https://doi.org/10.1177/0146167214552789>
- Shalvi, S., Gino, F., Barkan, R., & Ayal, S. (2015). Self-Serving Justifications: Doing Wrong and Feeling Moral. *Current Directions in Psychological Science*, 24(2), 125–130.
<https://doi.org/10.1177/0963721414553264>
- Skowronski, J. J., & Carlston, D. E. (1989). Negativity and extremity biases in impression formation: A review of explanations. *Psychological Bulletin*, 105(1), 131–142.
<https://doi.org/10.1037/0033-2909.105.1.131>
- Stanley, M. L., Henne, P., & De Brigard, F. (2019). 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>
- 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.
<https://doi.org/10.1177/1745691616689495>
- Strohminger, N., & Nichols, S. (2014). The essential moral self. *Cognition*, 131(1), 159–171.
- Vonasch, A. J., Reynolds, T., Winegard, B. M., & Baumeister, R. F. (2018). Death before dishonor: Incurring costs to protect moral reputation. *Social Psychological and Personality Science*, 9(5), 604–613. <https://doi.org/10.1177/1948550617720271>
- Weiner, B., Graham, S., Peter, O., & Zmuidinas, M. (1991). Public confession and forgiveness. *Journal of Personality*, 59(2), 281–312. <https://doi.org/10.1111/j.1467-6494.1991.tb00777.x>
- Wilson, A., & Ross, M. (2003). The identity function of autobiographical memory: Time is on our side. *Memory*, 11(2), 137–149.
- Wohl, M. J. A., Kim, H. S., Salmon, M., Santesso, D., Wildschut, T., & Sedikides, C. (2018). Discontinuity-induced nostalgia improves the odds of a self-reported quit attempt among people living with addiction. *Journal of Experimental Social Psychology*, 75, 83–94.
<https://doi.org/10.1016/j.jesp.2017.11.011>