

Unclearly Immoral: Low Self-Concept Clarity Increases Moral Disengagement

Jiaqian Wang^{1, 2} and Maferima Touré-Tillery³

¹ Department of Marketing, NUS Business School, National University of Singapore

² Global Asia Institute, National University of Singapore

³ Kellogg School of Management, Northwestern University

This research examines the effect of self-concept clarity (i.e., having self-beliefs that are clearly and confidently defined, internally consistent, and stable) on moral behavior. Seven preregistered studies ($N = 3,373$) document that low (vs. high) self-concept clarity decreases moral behavior (e.g., donation, volunteering, tax compliance, honesty in an incentivized game). This effect occurs because low self-concept clarity increases moral disengagement, leading people to behave in morally questionable manners without damaging their self-concept. As evidence for this proposed underlying mechanism, we show that the effect of self-concept clarity on moral behavior is mediated by state moral disengagement and attenuates (a) among people with low trait moral disengagement, (b) among people with high trait moral identity internalization, and (c) in the presence of an honor pledge cueing moral engagement. We then show that the effect holds only when a prosocial act is congruent with personal values. Overall, these findings contribute to the literature on self-concept and moral self-regulation and have implications for how to promote morality and curb unethical behavior in society.

Public Significance Statement

People continually face moral decisions, which collectively have a substantial impact on the functioning of their communities and of society at large. The present research reveals that lacking a clear sense of self (i.e., low self-concept clarity) decreases moral behavior (e.g., tax compliance, charitable giving, volunteering). This is because individuals with low self-concept clarity tend to be morally disengaged so that they can act in morally questionable ways without jeopardizing their self-concept. We show that making the moral implications of one's behaviors salient mitigates this effect. We also discuss potential interventions to help individuals maintain or increase their self-concept clarity. Overall, our work provides policymakers, employers, brands, and educators with strategies to curb moral disengagement and cultivate moral behavior.

Keywords: self-concept clarity, moral disengagement, morality, ethics, self

Supplemental materials: <https://doi.org/10.1037/xge0001654.supp>

You cannot transcend what you do not know. To go beyond yourself, you must know yourself.


—Sri Nisargadatta Maharaj

The self-concept is a collection of thoughts, ideas, and pieces of knowledge (e.g., traits, identities, values, goals) that people use to define who they are (Markus & Wurf, 1987). Self-concept clarity is the extent to which the self-concept is clearly and confidently defined, internally consistent, and stable over time (Campbell et al., 1996). The extant literature shows that low (vs. high) self-concept clarity can bring about a host of negative outcomes. For example,

relative to those high in self-concept clarity, people low in self-concept clarity tend to experience lower subjective well-being (Ritchie et al., 2011), greater daily negative affect (Nezlek & Plesko, 2001), higher anxiety and depression (Butzer & Kuiper, 2006), and less satisfying romantic (Lewandowski et al., 2010) and parent–child relationships (Becht et al., 2017).

Low self-concept clarity also impedes self-regulation, that is, the processes through which people influence, modify, or control their emotions, desires, and actions in the course of goal pursuit (Baumeister & Heatherton, 1996). People low (vs. high) in self-

This article was published Online First October 10, 2024.
Eddie Brummelman served as action editor.

Jiaqian Wang  <https://orcid.org/0000-0001-7762-1110>

An earlier version of this project was presented at the 15th Annual Meeting of the Society for the Science of Motivation on May 25, 2023, and the Association for Consumer Research Conference on October 27, 2023. Data, code, and preregistrations are available on the Open Science Framework (https://osf.io/bvhar/?view_only=26452037826c409589f6ce2dba142b04). Materials are available in Supplemental Materials.

Jiaqian Wang played a lead role in data curation, formal analysis, investigation,

methodology, resources, validation, visualization, and writing—original draft and an equal role in conceptualization, project administration, and writing—review and editing. Maferima Touré-Tillery played a lead role in funding acquisition and supervision, a supporting role in methodology, resources, validation, and writing—original draft, and an equal role in conceptualization, project administration, and writing—review and editing.

Correspondence concerning this article should be addressed to Jiaqian Wang, Department of Marketing, NUS Business School, National University of Singapore, BIZ 1, 8-13, Mochtar Riady Building, 15 Kent Ridge Drive, Singapore 119245, Singapore. Email: janewang@nus.edu.sg

concept clarity are less persistent and less diligent in pursuing their goals (Fite et al., 2017); for example, they spend less time practicing to enhance their task performance and are more likely to abandon their exercise plans (Jiang et al., 2023). Furthermore, low (vs. high) self-concept clarity is associated with emotional dysregulation (Parise et al., 2019), extravagant discretionary spending (Sarial-Abi et al., 2016), problematic gaming (Green et al., 2021), and compulsive internet and smartphone use (Quinones & Kakabadse, 2015; Servidio et al., 2021).

The present research extends this literature by exploring the consequences of low self-concept clarity on moral self-regulation (i.e., the processes through which individuals monitor and control their behaviors to adhere to moral standards, principles, and values; Bandura, 1991). We propose that low (vs. high) self-concept clarity impairs moral self-regulation by facilitating moral disengagement—that is, the dissociation between the self-concept and the moral implications of one's actions (Bandura et al., 1996). Specifically, we contend that people low (vs. high) in self-concept clarity are more likely to see their behaviors as justifiable and inconsequential, to perceive themselves as less responsible for their behaviors, and to lack empathy for those impacted by their behaviors. In turn, these perceptions allow people low (vs. high) in self-concept clarity to see morally questionable behaviors as less problematic, and thus less detrimental to their moral self-concept. Consequently, people low (vs. high) in self-concept clarity are less likely to act morally (e.g., decline to make a charitable donation) and more likely to act immorally (e.g., cheat on insurance or tax return).

By uncovering how self-concept clarity influences moral behavior, our work contributes to the literatures on self-concept clarity (Campbell, 1990; Campbell et al., 1996), moral self-regulation (Bandura et al., 1996), and the interplay between self-knowledge and moral behavior (Christy et al., 2016; Touré-Tillery & Light, 2018). Our findings also yield practical implications for fostering moral and prosocial behaviors and curbing unethical conduct.

Theoretical Development

Self-Concept Clarity and Moral Disengagement

Morality refers to a code of conduct advanced by an individual, group, or society regarding what is right and what is wrong (Gert & Gert, 2002). Moral behaviors (e.g., donating to an animal welfare charity) are those carried out according to moral values and standards (e.g., compassion for other living creatures), whereas immoral or unethical behaviors (e.g., cheating on an insurance claim) are characterized by deviations from moral principles (e.g., honesty). Often, being moral entails taking care of others' welfare even at the cost of one's self-interest, while acting immorally hurts others and may benefit oneself. For instance, volunteering requires time and effort that could have been spent on self-indulgent activities like a day trip or binge-watching a favorite Netflix series. Meanwhile, cheating on a car insurance claim may help one save or even earn more money at the expense of the insurance company and its stakeholders' (e.g., investors, employees) financial welfare.

Deciding whether to act immorally involves not only trade-offs between self-interest and the welfare of others but also considerations of how the behavior may affect one's self-concept. According to Bandura (1991), moral behaviors are guided by a moral self-regulation process, through which individuals monitor and regulate their actions based on how these actions may affect their moral self-

concept. Typically, moral behaviors boost one's moral self-concept (Gneezy, Imas, et al., 2012; Khan & Dhar, 2006), and moral transgressions damage one's sense of being a good, moral person (Smith et al., 2002). Because people are motivated to see themselves as good and moral (Steele, 1988; Strohminger & Nichols, 2014; Tesser, 1988), the moral self-regulatory process tends to guide them to behave morally and refrain from immoral acts (Bodner & Prelec, 2003; Gneezy, Gneezy, et al., 2012; Savary & Goldsmith, 2020).

We propose that low (vs. high) self-concept clarity impairs moral self-regulation by creating a dissociation between one's self-concept and the moral implications of one's actions—a cognitive process known as moral disengagement (Bandura et al., 1996). Greater moral disengagement means that for people low (vs. high) in self-concept clarity, their bad deeds have a lesser negative effect on their moral self-concept, such that they can derive the benefits of bad deeds at no cost to their self-concept.

Moral disengagement involves various interrelated cognitive mechanisms whereby individuals reframe their moral transgressions to make them seem more acceptable or less consequential or minimize their own moral agency in these transgressions. We posit that low (vs. high) self-concept clarity increases moral disengagement through these various mechanisms.

First, the self-concept plays an essential role in setting the standards that guide a person's behaviors (Adam et al., 2018; Higgins, 1996; Lewandowski & Nardone, 2012; Light, 2017). Thus, people high in self-concept clarity should have well-defined, consistent, and stable moral standards that they apply when making moral judgments and decisions. By contrast, people low in self-concept clarity should tend to have unclear, inconsistent, and unstable moral standards, such that a behavior deemed bad under one set of moral standards may be reframed as acceptable under a different set of standards. This reframing may involve portraying a morally questionable behavior as in the service of a more noble goal (i.e., moral justification; e.g., "I stole to help my friend;" see Thornberg & Jungert, 2014), comparing a morally questionable act to more reprehensible behaviors so that it seem benign or less consequential (i.e., advantageous comparison; e.g., "Stealing is nothing compared to killing someone;" see Corrion et al., 2009), or referring to a moral transgression in ways that mask its reprehensible nature (i.e., euphemistic labeling; e.g., "Civilians killed by the bombs are just 'collateral damage';" see Bandura, 1999).

Second, research shows that people low (vs. high) in self-concept clarity have trouble distinguishing between themselves and others. For example, they are more inclined to confuse their feelings, thoughts, and perspectives with those of others (Ickes et al., 2012). They are also more prone to "embody" a prosthetic hand or to have the illusion that another person's body is their own (Krol et al., 2020). Ironically, being able to separate the self from others is a critical prerequisite to empathy (i.e., the capacity to share the mental and emotional states of others) because this separation helps minimize the personal distress caused by an empathy-inducing situation (e.g., the suffering of others) and thus increases one's ability to cope with the situation. Further, self-other distinction helps reduce egocentric biases in assessing another person's circumstances (Batson et al., 1987; Lamm et al., 2016). Thus, people low (vs. high) in self-concept clarity tend to have less empathy (Krol & Bartz, 2022), which renders them less sensitive to the moral nature of a situation and less able to relate to how others are affected by their behaviors (Detert et al., 2008; Ogunfowora et al., 2022). We

propose that due to their lower empathy, people low (vs. high) in self-concept clarity are more likely to minimize the negative consequences of a moral transgression (i.e., distortion of consequences; e.g., “Teasing someone does not really hurt them;” Corrion et al., 2009), dehumanize the victims of their moral transgressions (e.g., “Some people deserve to be treated like animals;” Bandura et al., 1975), or even blame these victims as a way to justify a morally questionable behavior (e.g., “Kids who get mistreated usually do things that deserve it;” Thornberg & Jungert, 2014).

Finally, people low (vs. high) in self-concept clarity are more susceptible to external influences (Wang & Yu, 2023). For example, Mittal (2015) found that consumers low (vs. high) in self-concept clarity pay more heed to product recommendations when making purchases. Similarly, in indicating their personal beliefs, individuals low (vs. high) in self-concept clarity are more likely to align their answers with the consensus among their peers (Rahimi & Strube, 2007). Importantly, people lower in self-concept clarity are aware of their own receptivity to social influence, such that they are more inclined to attribute their own behaviors to external factors than to their internal traits, motivations, and values (Organ, 1973). In the context of moral transgressions, we propose that this tendency of people low (vs. high) in self-concept clarity to make external attributions for their own behaviors reduces their sense of moral agency, such that they feel less responsible for failing to act morally or for behaving immorally. For example, a person low (vs. high) in self-concept clarity may be more likely to attribute their shoplifting to peer pressure (i.e., displacement of responsibility, e.g., “My friends dared me to do it;” Milgram, 1963) or to social norm (i.e., diffusion of responsibility; “Everyone does it;” Bandura et al., 1975) rather than to their own dishonesty.

In sum, we propose that low (vs. high) self-concept clarity increases moral disengagement through the interconnected mechanisms of moral justification, advantageous comparison, euphemistic labeling, distortion of consequences, victim blaming, victim dehumanization, displacement of responsibility, and diffusion of responsibility. Furthermore, a large body of literature has documented that moral disengagement discourages good behaviors such as charitable giving (Hardy et al., 2015) while increasing moral transgressions such as cheating (e.g., doping, plagiarism; Detert et al., 2008; Hodge et al., 2013), aggression (e.g., bullying, Thornberg & Jungert, 2014), workplace misconduct (e.g., discussing confidential company information with an unauthorized person; Moore et al., 2012), and unethical business practices (Bandura et al., 2000; Moose, 2008). Thus, we advance that low (vs. high) self-concept clarity decreases moral behavior because it increases moral disengagement. We hypothesize:

Hypothesis 1: Low (vs. high) self-concept clarity decreases moral behavior.

Hypothesis 2: The effect of low (vs. high) self-concept clarity on moral behavior occurs through an increase in moral disengagement (mediation).

Given the proposed underlying role of moral disengagement in the effect of self-concept clarity on moral behavior, we predict that this effect will attenuate when moral disengagement is low—either chronically or situationally. This proposition is consistent with prior

findings that low moral disengagement renders people more immune to factors that may breed immoral behavior (Moore, 2015). For instance, negative emotional states increase counterproductive workplace behaviors among employees with high trait moral disengagement, but not among those with low trait moral disengagement (Samnani et al., 2014). Stated formally, we predict that:

Hypothesis 3: Moral disengagement moderates the effect of self-concept clarity on moral behavior, such that this effect attenuates or disappears among people who are chronically or situationally low in moral disengagement (i.e., those who are highly morally engaged).

The Roles of Moral Identity Internalization and Value Congruence

Moral identity internalization is the degree to which moral traits are personally important and central to a person’s self-concept (Aquino & Reed, 2002). People high in moral identity internalization tend to have chronically accessible moral self-schemas and standards, and thus are lower in moral disengagement (Detert et al., 2008). Like low moral disengagement, high moral identity internalization has been shown to attenuate factors that facilitate moral misconduct. For example, high moral identity internalization eliminates the effect of violent video game exposure on cyberbullying (Teng et al., 2020) and mitigates the effect of creativity on unethical workplace behavior (e.g., using company services for personal use; Zheng et al., 2019). Within this perspective, we propose that the effect of self-concept clarity on moral behavior attenuates among people high in moral identity internalization:

Hypothesis 4: Moral identity internalization moderates the effect of self-concept clarity on moral behavior, such that the effect attenuates or disappears among people high in moral identity internalization.

Finally, considering that the moral self-regulatory process relies on one’s personal standards and values, we predict that the effect of self-concept clarity on a particular moral behavior is contingent on the extent to which the behavior is congruent with the moral values that the decision maker *personally* holds. Specifically, this effect will occur only for moral behaviors congruent with one’s personal values. Although there are moral values that are universally held (e.g., respecting others’ property, helping others, reciprocating; Curry et al., 2019), some moral values may have different priority levels across different communities, while others may be held only within specific communities (Ellemers & van den Bos, 2012; Rai & Fiske, 2011). In the United States, for example, people who identify as political liberals more strongly endorse the moral values of caring about others and promoting fairness and equality, whereas political conservatives tend to put more emphasis on values such as loyalty to one’s ingroup, respect for authority, and purity (Graham et al., 2009). Consider, for example, a behavior such as choosing whether to help a charity advocating for the rights of lesbian, gay, bisexual, transgender, and queer (LGBTQ) people. In general, helping LGBTQ people reflects a concern for equal rights and autonomy,

which is more congruent with liberal values. Thus, we expect the effect of self-concept clarity on moral behavior to occur for liberals asked to help a charity supporting LGBTQ rights (i.e., a value-congruent act for liberals). By contrast, although helping others is a universal value, most conservatives might judge LGBTQ lifestyles as a violation of their purity value (de Zavala et al., 2014), such that helping LGBTQ people would not be value-congruent (or would even be value-incongruent) for them. Therefore, we predict that the effect of self-concept clarity on moral behavior will not occur for conservatives asked to help a charity supporting the rights of LGBTQ people (i.e., a value-incongruent act for conservatives). More generally, we hypothesize:

Hypothesis 5: Value congruence moderates the effect of self-concept clarity on moral behavior, such that the effect occurs only for moral behaviors congruent with one's values.

Overview of Studies

We tested our hypotheses about the effect of self-concept clarity on moral behavior in seven studies using various research designs (multiwave, cross-sectional, correlational, and experimental), participant populations (United States and China), and operationalizations of our key constructs. The first set of studies examined the main effect of self-concept clarity on moral behavior (Hypothesis 1). Study 1, conducted among students at a U.S.-based university, tested whether trait self-concept clarity measured at Time 1 predicted students' intentions to act morally in different scenarios at Time 2 (3 months later). Using a Chinese sample and an incentive-compatible design, Study 2 tested whether participants low (vs. high) in trait self-concept clarity were less likely to donate their bonus compensation (for participating in the study) to a charity. To establish causality, Study 3 employed an incentivized coin flip task to test whether an experimental manipulation of low (vs. high) self-concept clarity increased cheating for monetary gain among U.S.-based adults.

The next set of studies tested the proposed underlying role of moral disengagement through mediation (Study 4) and moderation (Studies 5 and 6). Specifically, Study 4 examined the mediating role of moral disengagement in the effect of self-concept clarity on tax compliance intentions (Hypothesis 2). The study also explored an important alternative mechanism: self-esteem. Study 5 investigated the moderating roles of trait moral disengagement (Hypothesis 3) and trait moral identity internalization (Hypothesis 4). To increase the external validity of our findings, this study manipulated self-concept clarity through a message embedded in a car insurance advertisement and tested how the message affected participants' likelihood of underreporting mileage to save money on the insurance premium. The study also examined the role of affect as an alternative process explanation. In Study 6, we tested the effectiveness of a theoretically and practically relevant intervention to curb dishonesty among people lacking self-concept clarity: cueing moral engagement through an honor pledge. Finally, Study 7 tested an important boundary condition for the effect of self-concept clarity on moral behavior: value congruence (Hypothesis 5). Specifically, the study explored how liberal versus conservative political views (a proxy for value congruence) moderate the effect of self-concept clarity on the decision to volunteer to support transgender rights.

Transparency and Openness

We report all sample size determinations, data exclusions (if any), experimental manipulations, and measures in each study. For each study, we aimed to recruit the preregistered number of observations, but the final number of completed responses may deviate slightly from the targeted number due to a glitch in the survey setup (Study 3) and the imprecise nature of participant counts on survey platforms (all studies). Full scales and materials are provided in [Supplemental Materials](#). Except for the Time 2 measures in Study 5, all studies were preregistered before data collection. Data, code, and preregistrations for all studies are available on the Open Science Framework (https://osf.io/bvhar/?view_only=26452037826c409589f6ce2dba142b04).

Study 1: Moral Scenarios

We conducted a multiwave survey study to test the link between trait self-concept clarity and moral decision making (Hypothesis 1). At Time 1, we administered Campbell et al.'s (1996) 12-item Self-Concept Clarity Scale, and at Time 2 (3 months later), participants completed a series of moral decision scenarios (Touré-Tillery & Light, 2018). We predicted a positive correlation between self-concept clarity and moral decision making. Put differently, we expected participants with lower self-concept clarity to be less likely to act morally.

Method

Participants

G*Power suggested that we needed at least 82 participants to detect a small-to-medium correlation of .30 with 80% power. We conducted data collection in two phases using the online student pool at a large U.S.-based university. All participants received monetary compensation. At Time 1, we recruited 246 participants ($M_{\text{age}} (N = 245) = 20.58$, $SD = 1.93$; 73 males, 161 females, 11 nonbinary). Three months later, we administered the Time-2 survey. We received responses from 149 returning participants (61% retention rate) within the preregistered 1-week window. All participants passed the preregistered attention check ("Select 'I'm not sure' for this question to show that you pay attention.") and remained in the final sample ($M_{\text{age}} = 20.87$, $SD = 2.01$; Gender: 46 males, 98 females, five nonbinary; race/ethnicity: 61 White, six Black or African American, 61 Asian, nine Hispanic or Spanish or Latino, one other, 11 Multiracial).

Procedure

At Time 1, participants completed the Campbell et al.'s (1996) 12-item Self-Concept Clarity Scale (e.g., "In general, I have a clear sense of who I am and what I am;" "My beliefs about myself often conflict with one another [reverse-coded];" 1 = *strongly disagree*, 7 = *strongly agree*) among other scales unrelated to the present investigation. We coded and averaged the items such that a higher score indicated stronger self-concept clarity ($\alpha = .87$; $M = 3.82$, $SD = 1.07$).

Three months later, participants who completed the Time-1 survey were invited to participate in the Time-2 survey, which

assessed the likelihood of engaging in 12 moral or immoral behaviors (e.g., “You are an intern at a company. Your coworker gets blamed for something you did. No one knows that you did it; not even your coworker. Would you confess that it was you and not your coworker?”; 1 = *absolutely not*, 5 = *absolutely yes*; see [Supplemental Materials](#) for full items; adapted from [Touré-Tillery & Light, 2018](#)). We coded and averaged the items such that a higher score indicated a greater likelihood to act morally ($\alpha = .69$; $M = 3.21$, $SD = 0.52$). We confirmed that participants who returned for the Time-2 survey did not differ in self-concept clarity from those who dropped out ($p > .47$, Cohen’s $d = .09$).

Results and Discussion

We found that participants with lower self-concept clarity were less likely to engage in moral behaviors ($r = .30$, $p < .001$; [Figure 1](#))—a result consistent with our hypothesis about the effect of self-concept clarity on moral behavior (Hypothesis 1).

Study 2: Donation Decisions

This study provided another test of the relationship between trait self-concept clarity and moral behavior (Hypothesis 1). To test the robustness and generalizability of our findings, the study (a) employed an incentive-compatible measure of moral behavior (small monetary donations) and (b) moved to a different cultural context (China). Research shows that Western and Eastern cultures differ in self-concept clarity and in the relationships between self-concept clarity and other constructs such as self-esteem and well-being (e.g., [Campbell et al., 1996](#)). Thus, this study examined whether our hypothesis generalizes beyond U.S.-based (Western) samples. After completing [Campbell et al.’s \(1996\)](#) 12-item Self-

Concept Clarity Scale, participants read a brief charitable appeal and decided whether to donate a small part of their compensation. We predicted a positive relationship between self-concept clarity and choosing to donate.

Method

Participants

We estimated that around 75% of participants would donate their bonus based on our experience with the subject pool. G*Power suggested that we needed at least 222 participants to detect a small effect (odds ratio $[OR] = 1.68$) with 80% power. We recruited 300 Chinese participants from Credamo and paid them for their time. As preregistered, we excluded two participants who failed the attention check, “Please choose strongly agree here to show that you are paying attention.” The final sample size was 298 ($M_{age} = 32.10$, $SD = 8.82$; gender/sex: 86 males, 212 females).

Procedure

The study employed a cross-sectional design. Participants first completed the same 12-item Self-Concept Clarity Scale as in Study 1 ([Campbell et al., 1996](#); $\alpha = .94$; $M = 4.59$, $SD = 1.50$). Then, they saw a thank you note informing them that, as a token of appreciation, each participant would receive a ¥0.50 bonus on top of their base compensation. Participants were asked to indicate if they would like to “donate my bonus to World Wildlife Fund” or “keep my bonus to receive it as extra compensation.”

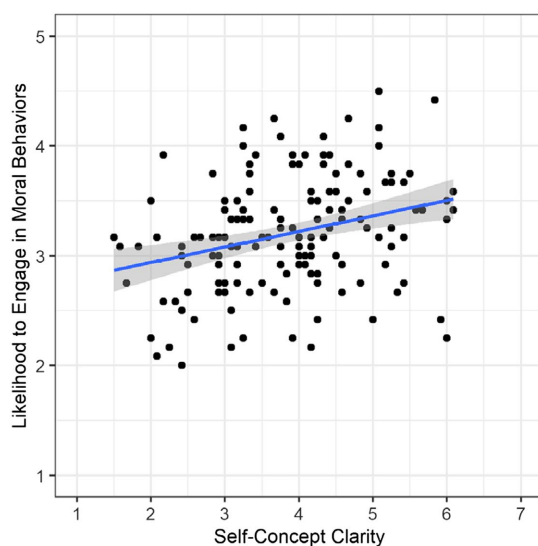
Results and Discussion

A logistic regression showed that participants with lower self-concept clarity were less likely to donate their bonus ($b = .47$, $SE = .09$, $z = 5.02$, $OR = 1.61$, $p < .001$; [Figure 2](#)). This study provided a conceptual replication of Study 1 using real donation decisions as our measure of moral behavior within a sample of Chinese adults. Taken together, the findings of Studies 1 and 2 demonstrate the robustness and the cultural generalizability of our hypothesis about the relationship between self-concept clarity and moral behavior (Hypothesis 1).

Study 3: Cheating to Earn an Extra Bonus

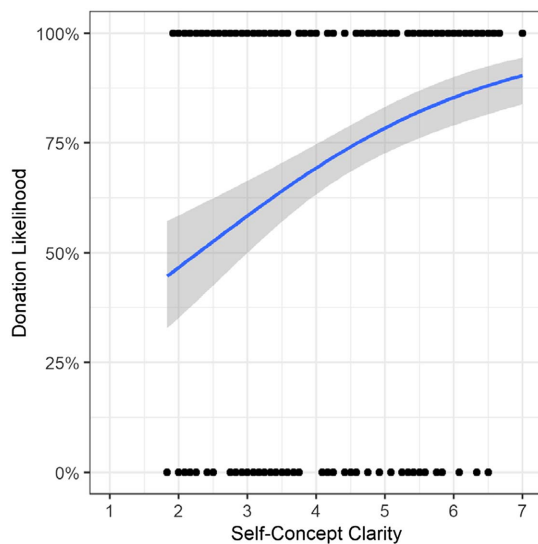
Although the results of the first two studies provided initial support for Hypothesis 1, the correlational designs of these studies do not allow us to ascertain the existence of a causal link between self-concept clarity and moral behavior. The present study addresses this limitation by testing the causal effect of self-concept clarity on moral behavior (Hypothesis 1) using an experimental design. After completing an experimental manipulation for self-concept clarity ([Wang & Yu, 2023](#)), participants took part in a study—ostensibly on probabilities—in which they were tasked with reporting the outcomes of five private coin flips. Half of the participants learned that they would receive additional compensation if the coin landed on heads (incentive condition), whereas the other half did not have this incentive (no-incentive condition). We intended for the monetary incentive and the privacy of the coin flip to offer a tempting opportunity for participants in the incentive condition to dishonestly

Figure 1
Likelihood to Engage in Moral Behaviors as a Function of Self-Concept Clarity (Study 1)



Note. See the online article for the color version of this figure.

Figure 2
Donation Likelihood as a Function of Self-Concept Clarity (Study 2)



Note. See the online article for the color version of this figure.

report more heads than they got. We could not determine whether the coins used by participants in this online study were fair, so the no-incentive condition served as a baseline level for coin flip outcomes. We predicted that incentivized participants in the low (vs. high) self-concept clarity condition would be more likely to (mis)report heads for the bonus compensation. However, we did not expect reports of heads to be different for nonincentivized participants in the low (vs. high) self-concept clarity condition. Furthermore, for participants in the low self-concept clarity condition, we predicted more reports of heads in the incentive (vs. no-incentive) condition—an indication of dishonest reporting. By contrast, for participants in the high self-concept clarity condition, we expected reports of heads to be similar regardless of the incentive structure—an indication of honest reporting.

Method

Participants

We predetermined a sample size of at least 130 valid responses per condition to detect a small-to-medium effect ($d = .35$) of self-concept clarity on cheating at 80% power. To allow for potential exclusions, we targeted a sample size of 650 Cloud Research-approved U.S.-based participants from Amazon's Mechanical Turk. Due to a small glitch in the survey setup at the initial stage of data collection, we received 663 responses. All participants received monetary compensation. As preregistered, we excluded 25 participants who provided invalid responses to the written task that manipulated self-concept clarity. The final sample size was 638 ($M_{\text{age}} = 39.87$, $SD = 12.51$; gender: 244 males, 387 females, seven nonbinary; race/ethnicity: 473 White, 60 Black or African American, one American Indian or Alaska Native, 34 Asian, 30 Hispanic or Spanish or Latino, five other, 35 Multiracial).

Procedure

The study used a 2 (Self-Concept Clarity: high vs. low) \times 2 (Financial Incentive: yes vs. no) between-subjects design. The first part of the experiment, presented as a "Life Experiences Survey," was a writing task that manipulated self-concept clarity using a procedure adopted from Wang and Yu (2023). Participants read six sentences adapted from the Self-Concept Clarity Scale items (Campbell et al., 1996) that, depending on the condition, described experiences evoking either high self-concept clarity (e.g., "My beliefs about myself are coherent with one another;" "My beliefs about myself don't change frequently") or low self-concept clarity (e.g., "My beliefs about myself conflict with one another;" "My beliefs about myself change frequently"). Participants were then asked to select a statement that aligned most closely with their experience and spend at least 1 min writing about that experience. This writing task was followed by a manipulation check: "Does the experience you just wrote about make you feel like you are very clear about who you are or like you are very unclear about who you are?" (1 = *very unclear*, 7 = *very clear*).

Next, participants proceeded to an ostensibly separate "Coin Flip Study," which constituted our measure of moral behavior using a procedure adapted from Touré-Tillery and Light (2018). The stated purpose of the study was to "understand the probabilities of different U.S. coins landing on heads or tails." In the incentive condition, participants learned that "As a token of appreciation, you will earn 5 cents every time the coin lands on heads." In the no-incentive condition, this information was omitted. We took several steps to ensure that all participants had a coin in hand. Those who reported that they did not have a coin with them right now were directed to a screen asking them to: "Please go get a U.S. coin now. Any denomination will work for this task. We will wait for you." Finally, once participants confirmed that they had a coin in hand, they reported the type of U.S. coin they had (e.g., 1 cent, 5 cents).

On the next page, participants were instructed to flip the coin five times and indicate whether the coin landed on heads or tails each time. We measured moral behavior through the number of heads reported, with larger numbers indicating dishonest reporting. Our rationale was that, because we could not observe the outcome of their coin flips, the incentive condition provided participants with a tempting opportunity to cheat by reporting more heads. By contrast, participants had no reason to be dishonest in the no-incentive condition, so this condition provided a baseline level of coin flip outcomes. This baseline was important because we had no control over the types of coins used by participants, such that we could not be sure that they would all be fair coins. Thus, we could not rely on 50% (i.e., the probability of getting heads with a fair coin) as the baseline of honesty. Instead, our design allowed us to infer dishonest behavior if incentivized participants reported more heads than nonincentivized participants.

Results and Discussion

Manipulation Check

We first checked the effectiveness of the self-concept clarity manipulation. Participants in the high (vs. low) self-concept clarity condition indeed indicated having clearer ideas about who they are

($M_{\text{high}} = 6.13, SD = 1.20$ vs. $M_{\text{low}} = 3.91, SD = 1.75$; $t = 18.64, p < .001, d = 1.48$).

Number of Heads Reported

A two-way analysis of variance of the number of heads reported on Self-Concept Clarity \times Financial Incentive revealed no main effects of self-concept clarity, $F(1, 634) = 1.04, p = .308, \eta_p^2 = .002$, or financial incentive, $F(1, 634) = 2.30, p = .130, \eta_p^2 = .004$. However, the predicted interaction of Self-Concept Clarity \times Financial Incentive emerged, $F(1, 634) = 5.70, p = .017, \eta_p^2 = .009$; Figure 3.

When reporting heads was not incentivized, participants in the high ($M = 2.71, SD = 0.89$) and low ($M = 2.61, SD = 0.83$) self-concept clarity conditions reported similar numbers of heads, $F(1, 634) = .96, p = .329$. In contrast, when reporting heads was incentivized, participants in the low self-concept clarity condition ($M = 2.89, SD = 0.96$) reported more heads than did those in the high self-concept clarity condition ($M = 2.65, SD = 0.95$; $F(1, 634) = 5.70, p = .017, d = .26$ [we report Cohen's d as the effect size in pairwise comparison for ease of comparison across studies]). Moreover, in the high self-concept clarity condition, the financial incentive did not affect the number of heads reported, $F(1, 634) = .38, p = .539$, whereas participants in the low self-concept clarity condition reported more heads in the presence (vs. absence) of a financial incentive, $F(1, 634) = 7.62, p = .006, d = .31$.

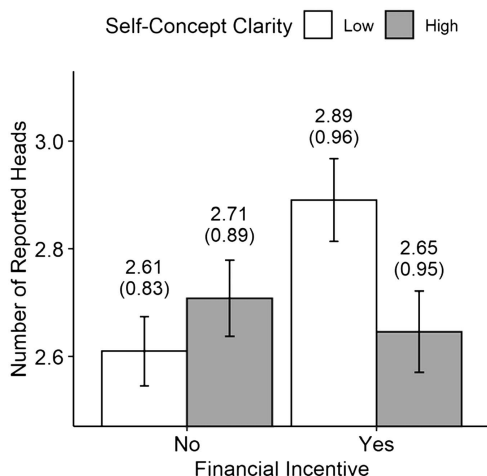
These results indicated that participants low (vs. high) in self-concept clarity were more likely to cheat for financial gain and provided causal evidence for our hypothesis about the negative effect of low (vs. high) self-concept clarity on moral behavior (Hypothesis 1).

Study 4: Mediation by Moral Disengagement

We argue that the effect of self-concept clarity on moral behavior documented in Studies 1 through 3 occurs because people low (vs.

Figure 3

Number of Reported Heads as a Function of Self-Concept Clarity and the Presence of a Financial Incentive (Study 3)



Note. Error bars indicate standard errors. Numbers above the bars indicate means (standard deviations).

high) in self-concept clarity have a greater tendency to morally disengage. Study 4 examines this proposed underlying mechanism by testing the mediating role of moral disengagement in the effect of self-concept clarity on moral behavior (Hypothesis 2) in the context of tax compliance. We predicted that lower trait self-concept clarity would be associated with lower tax compliance intentions and that this relationship would be mediated by moral disengagement.

In addition, we explored the role of self-esteem (i.e., the subjective, affect-laden appraisals of one's own value; Robins et al., 2001) as a potential alternative explanation for the effect of self-concept clarity on moral behavior. Indeed, research shows that self-concept clarity and self-esteem are positively correlated (Campbell, 1990). Furthermore, previous research has investigated the effect of self-esteem on moral behavior with mixed results: Self-esteem may increase (Aronson & Mettee, 1968), decrease (Baumeister et al., 1996), or have no effect on (Cheng, 2014) moral behavior. It is therefore possible that self-concept clarity increases moral behavior through an increase in self-esteem. Thus, in this study, we measured self-esteem to explore (a) its effect on tax compliance intentions and moral disengagement, (b) whether self-concept clarity predicts tax compliance intentions and moral disengagement when controlling for self-esteem (Lodi-Smith & DeMarree, 2017), and (c) whether self-esteem mediates the effect of self-concept clarity on tax compliance intentions.

Method

Participants

G*Power suggested that we needed at least 191 participants to detect a small-to-medium correlation of .20 with 80% power. We recruited 300 U.S.-based participants from Connect and paid them for their time. After preregistered exclusions based on attention checks ("Select 'strongly agree' for this question to show that you are paying attention." "Select 'strongly disagree' for this question to show that you are paying attention."), the final sample size was 292 ($M_{\text{age}} (N = 291) = 39.98, SD = 11.42$; gender: 146 males, 146 females; race/ethnicity: 226 White, 30 Black or African American, one American Indian or Alaska Native, 14 Asian, ten Hispanic or Spanish or Latino, two other, 9 Multiracial).

Procedure

The study employed a cross-sectional design. We first measured self-concept clarity using the same 12-item, 7-point scale as in Studies 1–2 ($\alpha = .95$; $M = 4.93, SD = 1.55$; Campbell et al., 1996).

Next, to the extent that tax compliance reflects people's willingness to give up personal gains for the public interest, we measured tax compliance intentions as a proxy for moral behavior. Following Achar and Lee's procedure (2022; Study 3), we asked participants to imagine that they would like to buy a car but could only afford it by underreporting their income to avoid taxes. Participants then indicated their likelihood to "Pay the income tax, even though it means I can't buy the car I want," "Underreport my income, so that I can buy the car I want (reverse-coded)," and "Not pay the income tax at all (reverse-coded)" on 10-point scales (1 = definitely no, 10 = definitely yes; $\alpha = .80, M = 8.31, SD = 2.09$).

To assess moral disengagement in the tax compliance context, we adapted eight items from established scales (Bandura et al., 1996;

Detert et al., 2008), such that each item captured one of the eight moral disengagement mechanisms (1 = *strongly disagree*, 7 = *strongly agree*): (a) euphemistic labeling: “Underreporting income is just ‘smart financial planning.’” (b) moral justification: “It is alright to avoid taxes to meet my own needs or my family’s needs.” (c) advantageous comparison: “Compared to other illegal activities such as robbery, avoiding taxes is not very serious.” (d) displacement of responsibility: “If someone is struggling to make ends meet, they cannot be blamed for avoiding taxes.” (e) diffusion of responsibility: “Someone cannot be blamed for avoiding taxes when many people do it.” (f) distortion of consequences: “It is okay to avoid some taxes because it doesn’t really do any harm.” (g) dehumanization: “Tax authorities are faceless entities.” and (h) attribution of blame: “People are not at fault for avoiding taxes if the government fails to address tax loopholes.” Principal component analysis with varimax rotation revealed that the eight items loaded on a single factor with an eigenvalue of 4.54 explaining 56.81% of the variance in the data. The eight items also showed good internal reliability ($\alpha = .88$). We thus averaged the items to form an index of moral disengagement ($M = 3.12$, $SD = 1.45$).

Finally, we assessed self-esteem using Robins et al.’s (2001) one-item self-esteem scale (i.e., “I have high self-esteem.” 1 = *not very true of me*, 7 = *very true of me*; $M = 4.66$, $SD = 1.81$).

Results and Discussion

Tax Compliance

We found a positive correlation between self-concept clarity and tax compliance intentions ($r = .26$, $p < .001$): as participants’ levels of self-concept clarity decreased, their intentions to dishonestly underpay taxes to buy the car they wanted increased.

Moral Disengagement

We observed a negative correlation between self-concept clarity and moral disengagement ($r = -.32$, $p < .001$): as participants’ levels of self-concept clarity decreased, they became more likely to morally disengage from tax compliance. We also note that self-concept clarity was negatively correlated with each of the eight moral disengagement items (see details in Supplemental Materials).

Mediation Analysis

A mediation analysis (PROCESS Model 4, Hayes, 2017) revealed a significant indirect effect of self-concept clarity on tax compliance intentions through moral disengagement ($b = .25$, Boot 95% CI [.15, .38]), such that participants with lower self-concept clarity had a greater tendency to morally disengage and in turn, were more inclined to dishonestly underpay taxes. Thus, Study 4 provided evidence for the mediating role of moral disengagement in the effect of self-concept clarity on moral behavior (Hypothesis 2).

Exploratory Analyses of the Effects of Self-Esteem

First, we examined bivariate correlations between self-esteem, tax compliance, and moral disengagement. We found that self-esteem was associated with lower tax compliance intentions ($r = -.14$, $p = .016$) but was uncorrelated with moral disengagement ($r = .03$, $p = .617$). Second, we tested whether self-concept clarity predicted tax

compliance and moral disengagement when controlling for self-esteem. A regression of tax compliance intentions on self-concept clarity and self-esteem showed divergent effects of the two predictors: While increasing self-concept clarity led to an increase in tax compliance intentions ($\beta = .36$, $t = 6.11$, $p < .001$), increasing self-esteem led to a decrease in tax compliance intentions ($\beta = -.27$, $t = -4.63$, $p < .001$). Another regression of moral disengagement on self-concept clarity and self-esteem revealed that increasing self-concept clarity decreased moral disengagement ($\beta = -.39$, $t = -6.54$, $p < .001$), whereas increasing self-esteem increased moral disengagement ($\beta = .17$, $t = 2.90$, $p = .004$). Finally, a mediation analysis (PROCESS Model 4, Hayes, 2017) with moral disengagement and self-esteem as parallel mediators showed a significant positive indirect effect through moral disengagement ($b = .24$, Boot 95% CI [.14, .36]) and a significant negative indirect effect through self-esteem ($b = -.09$, Boot 95% CI [-.16, -.03]). Overall, these results demonstrate the distinctness of self-concept clarity and self-esteem and the incremental predictive validity of self-concept clarity. Importantly, the findings show that the effects of self-concept clarity on moral disengagement and tax compliance are unlikely to stem from self-esteem—and in this study, these effects occur despite the opposing effect of self-esteem.

Study 5: Moderation by Trait Moral Disengagement and Trait Moral Identity Internalization

So far, four studies have provided convergent evidence that self-concept clarity influences moral behavior using different operationalizations of moral behavior (hypothetical decisions, small donations, cheating for extra compensation) and of self-concept clarity (measured through a scale, manipulated through a writing task). Although these operationalizations of self-concept clarity have high internal validity, they are not readily applicable to practical settings. Study 5 addresses this limitation by using a more ecologically valid manipulation of self-concept clarity to test the effect of self-concept clarity on moral behavior (Hypothesis 1) and the moderating roles of trait moral disengagement (Hypothesis 3) and trait moral identity internalization (Hypothesis 4).

Using a paradigm adapted from Wang and Yu (2023), we embedded a manipulation of self-concept clarity in an auto insurance advertisement that participants read at the start of the study. As a measure of moral behavior, participants then considered a scenario in which they could dishonestly underreport the mileage on their newly acquired car to secure a lower insurance premium. We also measured participants’ mood as a potential alternative mechanism, given that previous research has linked affect to both self-concept clarity (Nezlek & Plesko, 2001) and moral behavior (Vincent et al., 2013). Two weeks later, we invited the same participants to another survey wherein we measured their trait moral disengagement and moral identity internalization. We predicted that low (vs. high) self-concept clarity would increase the likelihood of dishonest reporting to the insurance company, but that this effect would not occur among people lower in moral disengagement (i.e., those who are more morally engaged) or among people higher in moral identity internalization.

Finally, as an additional test of our proposed mechanism, we measured trait perceptions of self-diagnosticsity as a possible moderator. Because trait self-diagnosticsity captures people’s general tendency to perceive their behaviors as reflective of

who they are (see Touré-Tillery & Light, 2018), high trait self-diagnosticity may create barriers to moral disengagement, thereby mitigating the effect of low (vs. high) self-concept clarity on moral behavior. Specifically, we tested the possibility that the effect of self-concept clarity on moral behavior would not occur at higher levels of trait self-diagnosticity.

Method

Participants

We predetermined a sample size of at least 130 valid responses per condition to detect a small-to-medium effect ($d = .35$) at 80% power. To allow for potential exclusions and given that the study used a relatively subtle manipulation of self-concept clarity, we first recruited 450 U.S.-based participants from Prolific. After exclusion based on preregistered attention checks, the final sample size was 440 ($M_{\text{age}} = 37.37$, $SD = 12.20$; gender: 219 males, 219 females, two nonbinary; race/ethnicity: 312 White, 43 Black or African American, one American Indian or Alaska Native, 43 Asian, 13 Hispanic or Spanish or Latino, three other, 25 Multiracial).

Two weeks later, we invited the same participants back for an ostensibly unrelated survey, in which we measured their chronic moral disengagement, moral identity internalization, and perceptions of self-diagnosticity. The survey was open for 72 hr. We received responses from 320 returning participants (72.7% retention rate) and paid them for their time. After excluding 11 participants who failed the attention check (i.e., "Select strongly agree here"), the final sample consisted of 309 responses for moderation analyses. All participants received monetary compensation.

Procedure

The study employed a two-condition (self-concept clarity: high vs. low) between-subjects design. Participants were asked to review an ad for "SuperCar," a fictional auto insurance company. We manipulated self-concept clarity through the content of the ad (a design adapted from Wang & Yu, 2023, Study 5). Specifically, the ad read:

[Low self-concept clarity] Life is a constantly changing journey. Think about those times that you act like very different persons in different situations, transition into a new position or place. ... All those experiences make you feel uncertain about yourself: "Is this really who I am?"

[High self-concept clarity] Life is a continuous journey. Think about those times that you are always the same person in different life roles, stay in the same place for many years. ... You say to yourself many times: "This is really who I am!" "I know myself so well!"

Let SuperCar join your journey! Introducing SuperCar. Easy, Affordable Auto Insurance.

As a manipulation check, participants answered four bipolar questions, "When thinking about the life experiences described in this ad, I feel like my beliefs and views about myself are" (1 = *very conflicting/uncertain/unclear/unstable*, 7 = *very coherent/certain/clear/stable*; $\alpha = .96$).

Then, we used a scenario adapted from Gai and Puntoni (2021) to measure dishonesty. Participants imagined that they had acquired a used car about a month earlier; now they wanted to purchase SuperCar insurance for their vehicle. SuperCar required them to

report the current odometer mileage to determine how much they would have to pay for the insurance. If they reported mileage lower than or equal to 5,000 miles, the annual premium would be \$700. If they reported mileage above 5,000 miles, they would have to pay \$1,200 or more per year. They further read that when they had purchased the car, its odometer had read 6,000 miles; now, after a month of driving, that figure had risen to 6,500 miles. Moreover, there was no way for SuperCar to verify the actual mileage; thus, the company would believe them if they said their car's mileage was 5,000 miles or lower. Participants were asked to answer the question "How many miles would you claim your car has to SuperCar?" Following Gai and Puntoni (2021), we coded any response equal to or below 5,000 miles as dishonest reporting and any response strictly above 5,000 miles as honest reporting. This question was followed by three attention checks regarding the scenario (see full questions in Supplemental Materials). Ten participants who failed all three checks were excluded. The study results did not change when we stringently excluded participants who failed any of the checks.

Finally, we used a subset of items from the Positive and Negative Affect Schedule (PANAS; Watson et al., 1988) to measure participants' positive (interested, enthusiastic, determined; $\alpha = .85$) and negative (upset, scared, afraid; $\alpha = .86$) affect (1 = *not at all*, 7 = *very much so*). The experiment ended with a demographic questionnaire and debriefing information.

Two weeks later, we invited the same participants back for an ostensibly unrelated survey, in which we measured their (a) trait moral disengagement (10 items; e.g., "Sometimes getting ahead of the curve is more important than adhering to rules"; $\alpha = .84$; Detert et al., 2008; Shu et al., 2011), (b) trait moral identity internalization (four items; e.g., "It would make me feel good to be an honest person"; $\alpha = .80$; adapted from Aquino & Reed, 2002), and (c) chronic perceptions of self-diagnosticity (seven items; e.g., "What I do is a reflection of who I am"; 1 = *strongly disagree*, 7 = *strongly agree*; $\alpha = .95$; Touré-Tillery & Light, 2018). The three scales appeared in a randomized order; and within each scale, the items were fully randomized.

Moral disengagement and moral identity internalization were negatively correlated ($r = -.44$, $p < .001$). Self-diagnosticity was positively correlated to moral identity internalization ($r = .34$, $p < .001$) and negatively correlated to moral disengagement ($r = -.24$, $p < .001$). None of the three measures differed across the self-concept clarity conditions randomly assigned 2 weeks prior ($ps > .29$).

Results and Discussion

Manipulation Check

Participants in the high (vs. low) self-concept clarity condition reported higher self-concept clarity ($M_{\text{high}} = 4.91$, $SD = 1.49$ vs. $M_{\text{low}} = 4.49$, $SD = 1.52$; $t = 2.92$, $p = .003$, $d = .28$), indicating that the manipulation produced the intended effect.

Dishonest Reporting

Participants in the low self-concept clarity condition were more inclined to report the wrong mileage (41.52%) to save on their premium than participants in the high self-concept clarity condition

(32.41%; $b = -.20$, $SE = .10$, $p = .048$)—indicating the former were more dishonest.

Affect

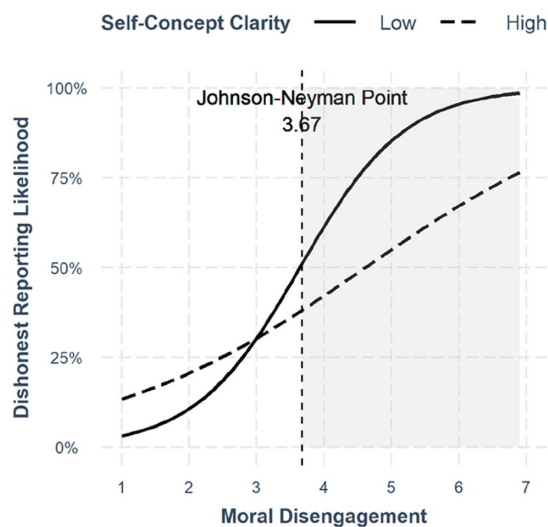
Self-concept clarity did not influence positive ($M_{\text{high}} = 4.01$, $SD = 1.53$ vs. $M_{\text{low}} = 3.94$, $SD = 1.49$; $t = .53$, $p = .595$, $d = .05$) or negative affect ($M_{\text{high}} = 1.61$, $SD = 0.98$ vs. $M_{\text{low}} = 1.66$, $SD = 1.04$; $t = .50$, $p = .616$, $d = .05$). Furthermore, positive affect did not influence dishonest reporting ($p > .42$), but greater negative affect increased participants' likelihood to cheat ($b = .32$, $SE = .10$, $p = .001$). Nonetheless, the null effects of self-concept clarity on affect suggested that affect is unlikely to explain the effect of self-concept clarity on dishonest reporting.

Moderation by Moral Disengagement ($M = 3.29$, $SD = 0.97$)

A logistic regression of dishonest reporting on self-concept clarity (high = 1, low = -1), mean-centered moral disengagement, and their interaction revealed no main effect of self-concept clarity ($b = -.11$, $SE = .13$, $p = .382$), but a main effect of moral disengagement ($b = .90$, $SE = .16$, $p < .001$) and a significant interaction of Self-Concept Clarity \times Moral Disengagement ($b = -.39$, $SE = .16$, $p = .012$; see Figure 4). Among more morally disengaged participants (Johnson-Neyman point of transition at 3.67; 33.98% of the sample), low (vs. high) self-concept clarity increased dishonest reporting, but this effect disappeared among participants lower in moral disengagement (i.e., the morally engaged, with a moral disengagement score below 3.67). Moreover, moral disengagement predicted dishonest reporting more strongly in the low self-concept clarity condition ($b = 1.29$, $SE = .24$, $p < .001$) than in the high self-concept clarity condition ($b = .52$, $SE = .20$, $p = .009$).

Figure 4

Dishonest Reporting Likelihood as a Function of Self-Concept Clarity and Moral Disengagement (Study 5)



Note. The vertical dotted line indicates the Johnson-Neyman point of transition. The shaded area indicates where the effect of self-concept clarity on dishonest reporting likelihood was statistically significant at $p < .05$. See the online article for the color version of this figure.

Moderation by Moral Identity Internalization ($M = 6.01$, $SD = 1.03$)

A regression of dishonest reporting on self-concept clarity, mean-centered moral identity internalization, and their interaction showed no main effect of self-concept clarity ($b = -.16$, $SE = .12$, $p = .202$), but a main effect of moral identity internalization ($b = -.52$, $SE = .12$, $p < .001$), and a marginally significant interaction of Self-Concept Clarity \times Moral Identity internalization emerged ($b = .21$, $SE = .12$, $p = .089$; see Figure 5). At lower levels of moral identity internalization (i.e., score below 5.48; 24.92% of the sample), participants in the low (vs. high) self-concept clarity condition were more likely to dishonestly report their mileage, but this effect did not occur at higher levels of moral identity internalization (score above 5.48). Moreover, the effect of moral identity internalization on cheating was stronger in the low self-concept clarity condition ($b = -.73$, $SE = .18$, $p < .001$) than in the high self-concept clarity condition ($b = -.31$, $SE = .17$, $p = .066$).

Moderation by Self-Diagnosticity ($M = 5.85$, $SD = 0.90$)

We regressed dishonest reporting on self-concept clarity, mean-centered trait self-diagnosticity, and their interaction term. We found no main effect of self-concept clarity ($b = -.12$, $SE = .12$, $p = .328$) and no interaction ($b = .08$, $SE = .13$, $p = .539$), but a negative main effect of self-diagnosticity on dishonest reporting emerged ($b = -.28$, $SE = .13$, $p = .033$): increasing perceptions of self-diagnosticity decreased dishonest reporting. We advance two possible (speculative) reasons for the absence of the predicted interaction. First, we note that the self-diagnosticity score in our sample was quite high and had a relatively low variance and a limited range (Median = 6.00 on a 7-point scale, range = [2.86, 7]; cf. Study 7 in Touré-Tillery & Wang, 2022, where trait self-diagnosticity was a significant moderator: $M = 5.33$, Median = 5.29, $SD = 0.98$, range = [1.43, 7]). These statistics suggest that we may have had too few participants and too little variance at lower levels of the self-diagnosticity scale to detect an effect of self-concept clarity. A second—though less likely—possibility is that the self-diagnosticity scale was not precise enough for our context. Indeed, the self-diagnosticity scale is designed to capture people's tendency to associate their *general* behaviors with their self-concept, whereas moral disengagement pertains to the tendency to dissociate *specific* types of behaviors (i.e., moral behavior) from their self-concept.

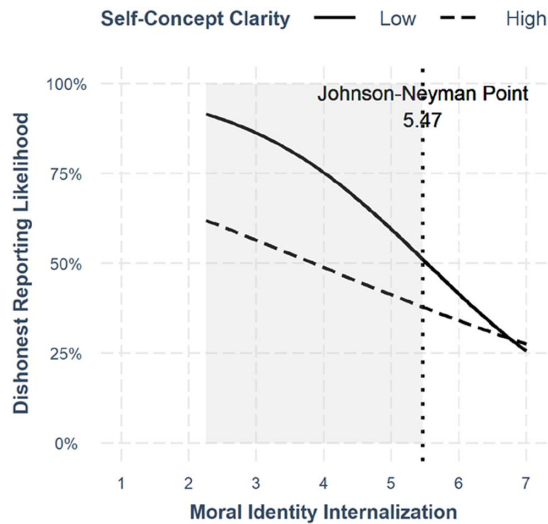
Using a manipulation of self-concept clarity that can be easily embedded in ads and other public messages, Study 5 demonstrated the robustness of the effect of low self-concept clarity on moral behavior (Hypothesis 1). Importantly, in line with the proposed underlying role of moral disengagement, the study showed that this effect disappeared among participants with lower trait moral disengagement (Hypothesis 3) or higher trait moral identity internalization (Hypothesis 4).

Study 6: Moderation by Moral Engagement

Study 6 examined whether cueing moral engagement through signing an honor pledge would mitigate the effect of low (vs. high) self-concept clarity on cheating for financial gain (Hypothesis 4). After completing the same manipulation of self-concept clarity as in Studies 3 and 4 (writing task), participants completed the

Figure 5

Dishonest Reporting Likelihood as a Function of Self-Concept Clarity and Moral Identity Internalization (Study 5)



Note. The vertical dotted line indicates the Johnson–Neyman point of transition. The shaded area indicates where the effect of self-concept clarity on dishonest reporting likelihood was statistically significant at $p < .05$. See the online article for the color version of this figure.

incentivized portion of the coin flip paradigm used in Study 3. As part of the instructions for this coin flip task, we randomly assigned half of the participants to sign an honor pledge that reminded them of what constitutes unethical behavior in the context of the task (moral engagement condition), whereas the other half saw no mention of an honor pledge (control condition). We expected to replicate the effect of self-concept clarity on moral behavior in the control condition, but not in the moral engagement condition.

Method

Participants

We predetermined a sample size of at least 130 valid responses per condition to detect a small-to-medium effect ($d = .35$) at 80% power. To allow for potential exclusions and to have enough power to detect an interactive effect, we aimed to recruit a sample size of 800. A total of 802 CloudResearch-approved U.S. participants completed this study in exchange for monetary compensation. After exclusions based on writing responses, the final sample size was 743 ($M_{\text{age}} = 39.60$, $SD = 12.12$; gender: 307 males, 423 females, 13 nonbinary; race/ethnicity: 544 White, 61 Black or African American, three American Indian or Alaska Native, one Native Hawaiian or Pacific Islander, 51 Asian, 32 Hispanic or Spanish or Latino, three other, 48 Multiracial).

Procedure

The study employed a 2 (Self-Concept Clarity: high vs. low) \times 2 (Moral Engagement: yes vs. control) between-subjects design. Participants first completed the same manipulation and manipulation check of self-concept clarity as in Studies 3 and 4, followed by the

incentivized portion of the coin flip task from Study 3. Participants flipped a coin five times and earned 5 cents for every “heads” they reported.

Right after the instructions for the coin flip task and before participants proceeded to flip their coin, we manipulated moral engagement as follows: Participant in the moral engagement condition saw an honor pledge intended to activate their moral self-regulation by providing statements explicitly describing what constitutes unethical behavior in the context of the coin flip task. The statements in the honor pledge were designed to counter moral disengagement mechanisms such as distorting the behavior, its consequences, and its victims. The honor pledge was as follows:

Before you proceed to flip the coin, please read each of the following statements and type your initials in the box next to each statement to affirm that you understand and agree

- it is unethical to lie to get a bonus in the study regardless of my personal circumstances
- it is wrong to steal even a small amount of money from a study
- cheating in the study is inappropriate because it hurts hardworking researchers who deserve to be treated with respect and honesty
- misrepresenting the truth in the study is bad because it hurts scientific advancement

Participants in the control condition did not see any mention of an honor pledge. As in Study 3, the number of heads reported served as our measure of dishonesty.

Results and Discussion

Manipulation Check

The manipulation of self-concept clarity was successful: Participants in the high (vs. low) self-concept clarity condition indicated having clearer ideas about who they were ($M_{\text{high}} = 6.02$, $SD = 1.21$ vs. $M_{\text{low}} = 4.07$, $SD = 1.78$; $t = 17.57$, $p < .001$, $d = 1.29$).

Number of Heads Reported

An analysis of variance of the number of heads reported on Self-Concept Clarity \times Moral Engagement showed no main effect of self-concept clarity, $F(1, 739) = 1.97$, $p = .161$, $\eta_p^2 = .003$, but a positive main effect of moral engagement, $F(1, 739) = 5.11$, $p = .024$, $\eta_p^2 = .007$, and a marginally significant interaction term, $F(1, 739) = 3.39$, $p = .066$, $\eta_p^2 = .005$; Figure 6. In the control condition, we replicated the results of Study 3 (for incentivized participants): The number of reported heads was higher in the low self-concept clarity condition ($M = 2.92$, $SD = 0.90$) than in the high self-concept clarity condition, $M = 2.69$, $SD = 0.92$; $F(1, 739) = .535$, $p = .021$, $d = .24$ —indicating more dishonesty in the former condition. However, in the moral engagement condition, the number of reported heads did not vary by self-concept clarity, $M_{\text{low}} = 2.63$, $SD = 1.02$ versus $M_{\text{high}} = 2.66$, $SD = 0.89$; $F(1, 739) = .09$, $p = .758$. From a different perspective, cueing moral engagement reduced the number of reported heads in the low self-concept clarity condition, $F(1, 739) = 8.35$, $p = .004$,

Figure 6
Number of Reported Heads as a Function of Self-Concept Clarity and Moral Engagement (Study 6)



Note. Error bars indicate standard errors. Numbers above the bars indicate means (standard deviations).

$d = .29$ but not in the high self-concept clarity condition, $F(1, 739) = .08$, $p = .767$.

In support of Hypothesis 4, we find that increasing moral engagement (using an honor pledge featuring explicit context-relevant examples of what constitutes unethical behaviors) mitigates the effect of self-concept clarity on cheating behavior. This honor pledge intervention is not only theoretically informative but also practically useful to curb the negative effect of low self-concept clarity on moral behavior—as it can be implemented easily in various contexts (e.g., work, school).

Study 7: Moderation by Value Congruence

In our studies so far, we have measured moral behavior using moral values that are relatively universal (e.g., honesty) and uncontroversial (e.g., caring for wildlife) in our participant populations. The broad endorsement of these values allowed us to document several main effects of self-concept clarity. However, there are contexts in which people's ideas about what is (im)moral may diverge, such that what is moral and good for one person may be amoral or even immoral for another person. For example, in the U.S. political context, supporting the proliferation of gender-neutral bathrooms to promote the well-being of transgender people is a polarizing cause, congruent with liberal ideologies but not with conservative ideologies (Prusaczyk & Hodson, 2020). To the extent that the moral self-regulatory process functions according to personal standards, the effect of self-concept clarity on a particular moral behavior should be contingent on the extent to which the behavior is congruent with the moral values that the decision maker *personally* holds (i.e., value-congruent; Hypothesis 5). In Study 7, we tested this boundary condition by examining the moderating role of value congruence in the context of volunteering to support the proliferation of gender-neutral bathrooms in the United States.

Participants completed the same writing manipulation of self-concept clarity as in previous studies. Then they indicated whether they would volunteer 2 min of their time to complete another survey to support transgender rights. At the end of the survey, we measured political views as a proxy for value congruence. We expected political views to moderate the effect of self-concept clarity on volunteering to promote gender-neutral bathrooms, such that low (vs. high) self-concept clarity would decrease such volunteering only among liberals (high value congruence)—but would not affect the decisions of conservatives (low value congruence).

Method

Participants

We predetermined a sample size of at least 130 valid responses per condition to detect a small-to-medium effect ($d = .35$) at 80% power. To allow for potential exclusions and to have enough power to detect an interactive effect, we aimed to recruit 900 U.S. participants from Prolific. For unknown reasons, we received complete responses from 891 participants and paid them for their time. After excluding participants who wrote gibberish in the self-concept clarity manipulation task or failed both attention checks (detailed later), we arrived at a final sample of 813 participants ($M_{age} = 37.54$, $SD = 13.57$; gender: 386 males, 411 females, 16 nonbinary; race/ethnicity: 607 White, 44 Black or African American, five American Indian or Alaska Native, 51 Asian, 51 Hispanic or Spanish or Latino, four other, 51 Multiracial).

Procedure

The study employed a 2 (Self-Concept Clarity: high vs. low) \times Political Views (measured) between-subjects design. Participants first completed the same writing task that constituted our manipulation of self-concept clarity in Studies 3 and 6. We then informed them that the researchers behind the study were supporting the efforts of the National Center for Transgender Equality (NCTE), an organization working to improve the well-being of transgender people. Participants read that the NCTE currently needs help to promote the establishment of gender-neutral bathrooms nationwide, and participants could help by completing an additional 2-min survey. We told them that they would receive no extra compensation for this second survey, but that the research team would donate \$0.25 to the NCTE on behalf of every participant who completed it. Participants indicated their volunteering decision by choosing “yes” or “no” in response to the question: “Would you like to help the NCTE by completing an extra 2-min survey?” Participants who decided to help proceeded to the second survey, which was just a filler survey. All participants then finished two attention checks: one on whether completing the survey could make extra money for themselves or a charity, and the other was a multiple-choice question on the charity's mission.

To check that political views indeed affected the extent to which promoting gender-neutral bathrooms was congruent with one's values, we assessed the perceived morality of promoting gender-neutral bathrooms (one item: 1 = *very immoral/bad*, 7 = *very moral/good*), alongside three filler issues (e.g., protecting gun rights).

Finally, after several demographic questions, we measured political views using four different measures in the interest of convergence: (a) political ideology (1 = *very liberal*, 7 = *very conservative*; reversed-coded to represent political liberalism), (b) voting history in the 2020 Presidential election (Biden, Trump, another candidate, did not vote), (c) political party affiliation (democrat, republican, other, not applicable), and (d) political identity (liberal, conservative, moderate, other, not applicable). We preregistered to conduct separate analyses with each of these four variables; for the three categorical variables, we planned to focus on the contrasts between Biden voters and Trump voters, democrats and republicans, and liberals and conservatives.

Results and Discussion

Manipulation Checks

The manipulation of self-concept clarity was successful, such that participants in the high (vs. low) self-concept clarity condition reported clearer self-views ($M_{\text{high}} = 6.03$, $SD = 1.20$ vs. $M_{\text{low}} = 3.83$, $SD = 1.78$; $t = 20.69$, $p < .001$, $d = 1.45$).

Moreover, political views indeed influenced the value congruence of promoting gender-neutral bathrooms. Political liberalism was positively correlated with the perceived morality of promoting gender-neutral bathrooms ($r = .73$, $p < .001$). Similarly, Biden (vs. Trump) voters ($M_{\text{Biden}} = 6.09$, $SD_{\text{Biden}} = 1.32$; $M_{\text{Trump}} = 2.60$, $SD_{\text{Trump}} = 1.67$; $t = 28.09$, $p < .001$, $d = 2.47$), democrats (vs. republicans; $M_{\text{Democrats}} = 6.09$, $SD_{\text{Democrats}} = 1.30$; $M_{\text{Republicans}} = 2.76$, $SD_{\text{Republicans}} = 1.75$; $t = 26.87$, $p < .001$, $d = 2.32$), and liberals (vs. conservatives; $M_{\text{liberals}} = 6.10$, $SD_{\text{liberals}} = 1.29$; $M_{\text{conservatives}} = 2.79$, $SD_{\text{conservatives}} = 1.75$; $t = 28.27$, $p < .001$, $d = 2.33$) perceived promoting gender-neutral bathrooms as more moral.

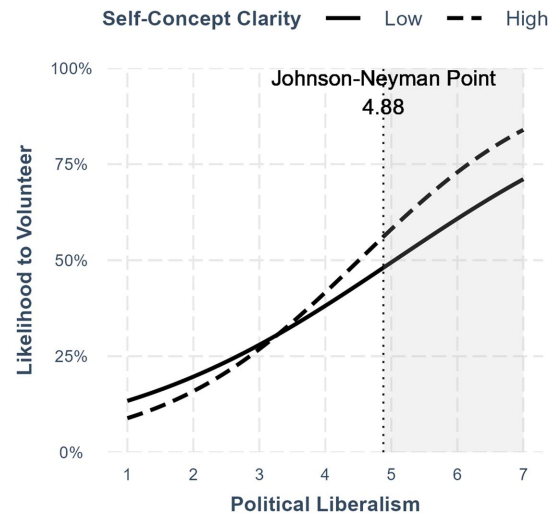
Volunteering Decision

A logistic regression of volunteering decision (yes = 1, no = 0) on self-concept clarity (high = 1, low = -1), mean-centered political liberalism, and their interaction term showed a positive main effect of self-concept clarity ($b = .19$, $SE = .08$, $p = .020$) and a positive main effect of political liberalism ($b = .56$, $SE = .05$, $p < .001$) on the likelihood to volunteer. Importantly, the interaction term was significant ($b = .10$, $SE = .05$, $p = .034$; Figure 7): High (vs. low) self-concept clarity increased volunteering likelihood among more liberal participants (i.e., participants who scored 4.88 or higher on the political liberalism scale; 72.57% of the sample), but not among more conservative participants (i.e., those with scores strictly lower than 4.88 on the political liberalism scale).

Similarly, there were significant interactions of self-concept clarity and (a) voting history (Biden vs. Trump; $b = -.52$, $SE = .23$, $p = .024$), (b) party affiliation (democrat vs. republican; $b = -.44$, $SE = .21$, $p = .034$), and (c) political identity (liberal vs. conservative; $b = -.43$, $SE = .20$, $p = .030$) in shaping the volunteering rate. Specifically, low (vs. high) self-concept clarity reduced the volunteering rate among Biden voters (61.83% vs. 76.36%; $b = .35$, $SE = .10$, $p < .001$), democrats (60.08% vs. 73.93%; $b = .32$, $SE = .10$, $p = .001$), and liberals (62.01% vs. 74.65%; $b = .30$, $SE = .09$, $p = .001$), but not among Trump voters (18.99% vs. 14.29%; $b = -.17$, $SE = .21$, $p = .411$), republicans

Figure 7

Likelihood to Volunteer to Promote Gender-Neutral Bathrooms as a Function of Self-Concept Clarity and Political Liberalism (Study 7).



Note. The vertical dotted line indicates the Johnson–Neyman point of transition. The shaded area indicates where the effect of self-concept clarity on likelihood to volunteer was statistically significant at $p < .05$. See the online article for the color version of this figure.

(22.99% vs. 18.95%; $b = -.12$, $SE = .18$, $p = .503$), or conservatives (22.22% vs. 17.82%; $b = -.14$, $SE = .18$, $p = .437$).

Study 7 provided evidence for the moderating role of value congruence (Hypothesis 5) by showing that low (vs. high) self-concept clarity decreased the likelihood of helping promote gender-neutral bathrooms among more liberal participants but had no effect among conservative ones.

General Discussion

The present research examines the effect of self-concept clarity on moral behavior. Across seven studies, using different operationalizations of self-concept clarity and both hypothetical (Studies 1, 4, and 5) and real/incentive-compatible (Studies 2, 3, 6, and 7) moral behaviors, we find that people low (vs. high) in self-concept clarity are less likely to behave morally. We show that this effect occurs through moral disengagement, such that it is statistically mediated by moral disengagement (Study 4), attenuates among people chronically low in moral disengagement or chronically high in moral identity internalization (Study 5), attenuates in the presence of an honor pledge cueing moral engagement (Study 6), and occurs only when the moral behavior is congruent with personal moral values (Study 7). Furthermore, our results suggest that the effect of self-concept clarity on moral behavior cannot be attributed to self-esteem (Study 4) or affect (Study 5). These findings have theoretical and practical implications on which we elaborate in the sections that follow.

Theoretical Implications

Our work contributes to several literatures. First, we extend research on the consequences of self-concept clarity. A large stream

of work on this topic has focused on the negative life outcomes of low self-concept clarity (e.g., poor mental health, low relationship satisfaction; Butzer & Kuiper, 2006; Nezlek & Plesko, 2001; Ritchie et al., 2011), as well as the ways in which people cope with the aversiveness of having low self-concept clarity (e.g., keeping identity-relevant subscriptions, choosing personality-revealing products; Rozenkrants et al., 2017; Savary & Dhar, 2020; Wang & Yu, 2023). Recent work has embarked on examining the self-regulatory consequences of low self-concept clarity (e.g., Fite et al., 2017; Jiang et al., 2023). These studies show that low self-concept clarity inhibits self-regulation in decisions involving trade-offs between present and future outcomes. For example, people low (vs. high) in self-concept clarity are more susceptible to discretionary spending (Sarial-Abi et al., 2016) and problematic gaming (Green et al., 2021), which brings immediate gratification at the expense of long-term pursuits and welfare. The *intertemporal* nature of these decisions is evidenced by the finding that reduced self-continuity—a weakened sense of connectedness among the past, present, and future selves—mediates the detrimental effect of low self-concept clarity on self-control (Jiang et al., 2023). In the current work, we show that low self-concept clarity leads to prioritizing self-interest over the welfare of others in moral decision making because it facilitates moral disengagement—which enables people to act immorally without damaging their self-concept. In doing so, we contribute to the burgeoning line of literature on how self-concept clarity affects self-regulation by highlighting that beyond *intrapersonal* preferences that affect long-term personal well-being, low self-concept clarity also influences *interpersonal* processes that affect societal functioning.

Further, our finding holds promise in providing an integrative perspective to understand the links between low self-concept clarity and self-regulatory failures. Within the realm of moral decision making, our work demonstrates that low self-concept clarity impairs moral self-regulation by facilitating the dissociation of the implications of one's moral behavior from one's self-concept. We view moral disengagement as a domain-specific perception that one's actions are not self-diagnostic. Extending this logic, we contend that low self-concept clarity may also reduce perceptions of self-diagnostics in general, such that people low (vs. high) in self-concept clarity may be less prone to see their behavior as indicative of who they are in domains outside of morality. Because perceptions of self-diagnostics are key drivers of self-regulation in various domains such as healthy eating and spending behaviors (Touré-Tillery & Fishbach, 2012, 2015; Touré-Tillery & Wang, 2022), decreased self-diagnostics may explain why low self-concept clarity impairs self-regulation in general (e.g., Fite et al., 2017; Jiang et al., 2023). We call for future research to test this possibility.

Moreover, given that low self-concept clarity is an aversive state, research shows that people low (vs. high) in self-concept clarity are more motivated to engage in behaviors that help them to achieve a clearer and more positive self-concept (Savary & Dhar, 2020; Wang & Yu, 2023). Thus, one might wonder why low (vs. high) self-concept clarity fails to similarly increase moral behavior which may provide a boost to the moral self-concept. We propose that this discrepancy may be due to two main factors: (1) whether people have clear guides in the given domain and (2) whether the behavior is perceived as instrumental in improving the self-concept.

Wang and Yu (2023) found that low (vs. high) self-concept clarity leads people to improve their physical attractiveness, a domain with

relatively socially constructed, homogeneous standards within a given society or community (Wolf, 2013). Furthermore, Savary and Dhar (2020) showed that consumers low (vs. high) in self-concept clarity are more likely to keep unused subscriptions that tell them something positive about themselves. These findings suggest that behaviors with clear standards and direct paths to self-concept clarification or enhancement are appealing to people with low self-concept clarity. In contrast, moral values are relatively idiosyncratic and diverse (Graham et al., 2009), which makes it challenging for people low (vs. high) in self-concept clarity to form a clear, coherent set of self-guides to navigate their moral decisions. Furthermore, our findings reveal that low self-concept clarity causes moral disengagement, diminishing the impact of (im)moral behavior on one's self-concept. Thus, people low (vs. high) in self-concept clarity may not only lack clear self-guides in moral decision making, but also see moral behavior as less instrumental for self-concept clarification or enhancement. Consequently, low self-concept clarity tends to reduce, rather than increase, engagement in moral behaviors.

It follows that the effect of self-concept clarity on moral behavior may be contingent on whether the decision is made in private or public settings. The present research documents that low (vs. high) self-concept clarity reduces moral behaviors in private contexts, which are presumably guided by internal self-guides and are not subject to social evaluations. To the extent that low (vs. high) self-concept clarity increases public self-consciousness and susceptibility to social influences (Wang & Yu, 2023), chances are that people low in self-concept clarity would act more morally or as morally as people high in self-concept clarity in public settings to appear moral in others' eyes, especially in environments that encourage moral behavior. It would be meaningful for future research to systematically test how self-concept clarity influences private versus public moral behavior, such as the decision to donate through an anonymous online platform versus at a social fundraising event.

Our findings also contribute to research on the antecedents of moral disengagement. Moral disengagement—referring to the dissociation of the *moral* implications of one's behavior from one's *self-concept*—presumably could be enabled by individuals' traits or experiences related to morality, or by characteristics of their self-concept. In examining precursors to moral disengagement, the extant literature has prominently focused on morality-related individual differences (e.g., weak moral identity, low honesty-humility, low moral idealism, high moral relativism, dark triad; Egan et al., 2015; Lian et al., 2022; Ogunfowora et al., 2022), situational factors (e.g., organizational injustice, unethical leadership; Hystad et al., 2014), and prior moral transgressions (e.g., Shu et al., 2011). By showing that low self-concept clarity increases moral disengagement, our research opens an avenue to examine how the self-concept may facilitate such disengagement. Future research could extend this line of investigation by testing how other aspects of the self (e.g., self-construal, self-complexity, self-compartmentalization) may affect moral disengagement.

In addition, our findings contribute to a deeper understanding of the interplay between self-knowledge and morality. Previous literature has documented that morality is crucial in maintaining a clear self-concept, such that people feel they know themselves less well on days when they are less moral or after recalling their past unethical acts (Christy et al., 2016). Our work reveals the bidirectional relationship between self-knowledge and morality:

knowing oneself well increases moral behaviors while curbing immoral ones. In doing so, we highlight that self-knowledge is not merely a passive reflection of one's moral standing but an active contributor to moral behavior.

Practical Implications

People today face various challenges in knowing themselves. The COVID-19 pandemic and accompanying economic and political climates have resulted in many factors known to disrupt a person's self-concept clarity (e.g., the passing of a loved one, unemployment, a lack of shared reality; Boelen et al., 2012; Light & Goldberg, 2020; McIntyre et al., 2014). Today's digitalized society provides people with unprecedented opportunities to try out distinct possible selves and receive a wide variety of feedback in online gaming or on social media, which may unintentionally create more confusion around the self-concept (Appel et al., 2018). Our findings suggest that unethical behavior may proliferate in such a context. Thus, more than ever, it is important to implement interventions that help maintain or increase self-concept clarity in society. For example, companies may conduct workshops or training sessions that guide employees in aligning their personal values with organizational values to contribute to a more coherent self-concept. It would also be helpful to foster a culture of open and inclusive communication so that employees feel comfortable expressing their ideas and values to reinforce who they are. Moreover, schools may incorporate regular self-reflection exercises into the curriculum or organize self-reflection activities (e.g., journaling) to help students reflect on their experiences, thoughts, emotions, strengths, weaknesses, goals, and values to have a deeper understanding of themselves. These interventions, to the extent that they increase or help to maintain self-concept clarity, would in turn promote ethical behavior.

Finally, our results underscore the importance of interventions that reduce moral disengagement to buffer against the negative consequences of self-concept clarity and increase moral behavior in general. Such interventions may tackle different mechanisms of moral disengagement. For example, people often justify lying on insurance claims or tax returns by saying that "Cheating the government or a big company of a few hundred bucks is not a big deal" (i.e., distortion of consequences) or "There are people who cheat more than I do" (i.e., advantageous comparison). To address this form of moral disengagement, it would be useful for institutions to make the case that cheating at any level is cheating, unfair, and harmful to fellow citizens. Moreover, in some cases, people justify treating others with less moral concern and empathy by denying others' human attributes such as their agency or ability to experience emotions (i.e., dehumanization). Nonprofit organizations may address this issue by spotlighting how people from marginalized or stigmatized groups feel about their suffering and strive to overcome their hardships.

Constraints on Generality

First, while our studies involved American and Chinese participants, the broader applicability and generalizability of our findings across diverse countries and cultures would necessitate further investigation. Second, we operationalized moral behavior through conduct that reflects honesty, compassion, generosity, and support for equity. Future research may examine whether the findings

are generalizable to alternative moral values such as respect for authority. Third, the present research tested moral behavior in mostly private settings; it is possible that self-concept clarity affects moral behavior differently in public settings. Examining this possibility could be a fruitful avenue for future research. Finally, in testing the moderating role of value congruence, Study 7 utilized American participants' political ideology as a proxy for their attitudes toward gender-neutral bathrooms. This decision was based on the political polarization surrounding LGBTQ issues at the time of the study in 2022, as well as the consideration that political ideology may be a less intrusive measurement than explicit, LGBTQ-specific attitude measurements. It is important to acknowledge that the correlation between political ideology and attitudes toward gender-neutral bathrooms is subject to dynamic shifts over time.

Coda

Sri Nisargadatta Maharaj said, "You cannot transcend what you do not know. To go beyond yourself, you must know yourself." The current research underscores this idea by revealing that not knowing oneself well prevents people from transcending their self-interest for the greater good. Put another way, having a clear sense of self lays the foundation for behaving in a clearly moral manner.

References

- Achar, C., & Lee, A. Y. (2022). Regulatory fit intensifies moral predispositions. *Journal of Personality and Social Psychology*, 123(3), 481–502. <https://doi.org/10.1037/pspa0000306>
- Adam, H., Obodaru, O., Lu, J. G., Maddux, W. W., & Galinsky, A. D. (2018). The shortest path to oneself leads around the world: Living abroad increases self-concept clarity. *Organizational Behavior and Human Decision Processes*, 145, 16–29. <https://doi.org/10.1016/j.obhdp.2018.01.002>
- Appel, M., Schreiner, C., Weber, S., Mara, M., & Gnambs, T. (2018). Intensity of Facebook use is associated with lower self-concept clarity. *Journal of Media Psychology: Theories, Methods, and Applications*, 30(3), 160–172. <https://doi.org/10.1027/1864-1105/a000192>
- 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>
- Aronson, E., & Mettee, D. R. (1968). Dishonest behavior as a function of differential levels of induced self-esteem. *Journal of Personality and Social Psychology*, 9(2), 121–127. <https://doi.org/10.1037/h0025853>
- Bandura, A. (1991). Social cognitive theory of self-regulation. *Organizational Behavior and Human Decision Processes*, 50(2), 248–287. [https://doi.org/10.1016/0749-5978\(91\)90022-L](https://doi.org/10.1016/0749-5978(91)90022-L)
- 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
- 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>
- Bandura, A., Caprara, G. V., & Zsolnai, L. (2000). Corporate transgressions through moral disengagement. *Journal of Human Values*, 6(1), 57–64. <https://doi.org/10.1177/097168580000600106>
- Bandura, A., Underwood, B., & Fromson, M. E. (1975). Disinhibition of aggression through diffusion of responsibility and dehumanization of victims. *Journal of Research in Personality*, 9(4), 253–269. [https://doi.org/10.1016/0092-6566\(75\)90001-X](https://doi.org/10.1016/0092-6566(75)90001-X)
- Batson, C. D., Fultz, J., & Schoenrade, P. A. (1987). Distress and empathy: Two qualitatively distinct vicarious emotions with different motivational

- consequences. *Journal of Personality*, 55(1), 19–39. <https://doi.org/10.1111/j.1467-6494.1987.tb00426.x>
- Baumeister, R. F., & Heatherton, T. F. (1996). Self-regulation failure: An overview. *Psychological Inquiry*, 7(1), 1–15. https://doi.org/10.1207/s15327965pli0701_1
- Baumeister, R. F., Smart, L., & Boden, J. M. (1996). Relation of threatened egotism to violence and aggression: The dark side of high self-esteem. *Psychological Review*, 103(1), 5–33. <https://doi.org/10.1037/0033-295X.103.1.5>
- Becht, A. I., Nelemans, S. A., van Dijk, M. P. A., Branje, S. J. T., Van Lier, P. A. C., Denissen, J. J. A., & Meeus, W. H. J. (2017). Clear self, better relationships: Adolescents' self-concept clarity and relationship quality with parents and peers across 5 years. *Child Development*, 88(6), 1823–1833. <https://doi.org/10.1111/cdev.12921>
- Bodner, R., & Prelec, D. (2003). Self-signaling and diagnostic utility in everyday decision making. In I. Brocas & J. D. Carrillo (Eds.), *The psychology of economic decisions: Rationality and well-being* (pp. 105–126). Oxford Academic. <https://doi.org/10.1093/oso/9780199251063.003.0006>
- Boelen, P. A., Keijsers, L., & van den Hout, M. A. (2012). The role of self-concept clarity in prolonged grief disorder. *Journal of Nervous and Mental Disease*, 200(1), 56–62. <https://doi.org/10.1097/NMD.0b013e31823e577f>
- Butzer, B., & Kuiper, N. A. (2006). Relationships between the frequency of social comparisons and self-concept clarity, intolerance of uncertainty, anxiety, and depression. *Personality and Individual Differences*, 41(1), 167–176. <https://doi.org/10.1016/j.paid.2005.12.017>
- Campbell, J. D. (1990). Self-esteem and clarity of the self-concept. *Journal of Personality and Social Psychology*, 59(3), 538–549. <https://doi.org/10.1037/0022-3514.59.3.538>
- Campbell, J. D., Trapnell, P. D., Heine, S. J., Katz, I. M., Lavallee, L. F., & Lehman, D. R. (1996). Self-concept clarity: Measurement, personality correlates, and cultural boundaries. *Journal of Personality and Social Psychology*, 70(1), 141–156. <https://doi.org/10.1037/0022-3514.70.1.141>
- Cheng, C. (2014). The predictive effects of self-esteem, moral self, and moral reasoning on delinquent behaviors of Hong Kong young people. *International Journal of Criminology and Sociology*, 3, 133–145. <https://doi.org/10.6000/1929-4409.2014.03.12>
- Christy, A. G., Seto, E., Schlegel, R. J., Vess, M., & Hicks, J. A. (2016). Straying from the righteous path and from ourselves: The interplay between perceptions of morality and self-knowledge. *Personality and Social Psychology Bulletin*, 42(11), 1538–1550. <https://doi.org/10.1177/0146167216665095>
- Corrion, K., Long, T., Smith, A. L., & d'Arripe-Longueville, F. (2009). "It's not my fault; it's not serious": Athlete accounts of moral disengagement in competitive sport. *The Sport Psychologist*, 23(3), 388–404. <https://doi.org/10.1123/tsp.23.3.388>
- Curry, O. S., Mullins, D. A., & Whitehouse, H. (2019). Is it good to cooperate? Testing the theory of morality-as-cooperation in 60 societies. *Current Anthropology*, 60(1), 47–69. <https://doi.org/10.1086/701478>
- de Zavala, A. G., Walduz, S., & Cyprianska, M. (2014). Prejudice towards gay men and a need for physical cleansing. *Journal of Experimental Social Psychology*, 54, 1–10. <https://doi.org/10.1016/j.jesp.2014.04.001>
- 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>
- Egan, V., Hughes, N., & Palmer, E. J. (2015). Moral disengagement, the dark triad, and unethical consumer attitudes. *Personality and Individual Differences*, 76, 123–128. <https://doi.org/10.1016/j.paid.2014.11.054>
- Ellemers, N., & van den Bos, K. (2012). Morality in groups: On the social-regulatory functions of right and wrong. *Social and Personality Psychology Compass*, 6(12), 878–889. <https://doi.org/10.1111/spc3.12001>
- Fite, R. E., Lindeman, M. I., Rogers, A. P., Voyles, E., & Durik, A. M. (2017). Knowing oneself and long-term goal pursuit: Relations among self-concept clarity, conscientiousness, and grit. *Personality and Individual Differences*, 108, 191–194. <https://doi.org/10.1016/j.paid.2016.12.008>
- Gai, P. J., & Puntoni, S. (2021). Language and consumer dishonesty: A self-diagnostics theory. *Journal of Consumer Research*, 48(2), 333–351. <https://doi.org/10.1093/jcr/ucab001>
- Gert, B., & Gert, J. (2002). The definition of morality. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of philosophy* (Fall 2020 ed.). The Metaphysics Research Lab, Center for the Study of Language and Information, Stanford University. <https://plato.stanford.edu/archives/fall2020/entries/morality-definition>
- Gneezy, A., Gneezy, U., Riener, G., & Nelson, L. D. (2012). Pay-what-you-want, identity, and self-signaling in markets. *Proceedings of the National Academy of Sciences of the United States of America*, 109(19), 7236–7240. <https://doi.org/10.1073/pnas.1120893109>
- Gneezy, A., Imas, A., Brown, A., Nelson, L. D., & Norton, M. I. (2012). Paying to be nice: Consistency and costly prosocial behavior. *Management Science*, 58(1), 179–187. <https://doi.org/10.1287/mnsc.1110.1437>
- Graham, J., Haidt, J., & Nosek, B. A. (2009). Liberals and conservatives rely on different sets of moral foundations. *Journal of Personality and Social Psychology*, 96(5), 1029–1046. <https://doi.org/10.1037/a0015141>
- Green, R., Delfabbro, P. H., & King, D. L. (2021). Avatar identification and problematic gaming: The role of self-concept clarity. *Addictive Behaviors*, 113, Article 106694. <https://doi.org/10.1016/j.addbeh.2020.106694>
- Hardy, S. A., Bean, D. S., & Olsen, J. A. (2015). Moral identity and adolescent prosocial and antisocial behaviors: Interactions with moral disengagement and self-regulation. *Journal of Youth and Adolescence*, 44(8), 1542–1554. <https://doi.org/10.1007/s10964-014-0172-1>
- Hayes, A. F. (2017). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. Guilford publications.
- Higgins, E. T. (1996). The "self digest": Self-knowledge serving self-regulatory functions. *Journal of Personality and Social Psychology*, 71(6), 1062–1083. <https://doi.org/10.1037/0022-3514.71.6.1062>
- Hodge, K., Hargreaves, E. A., Gerrard, D., & Lonsdale, C. (2013). Psychological mechanisms underlying doping attitudes in sport: Motivation and moral disengagement. *Journal of Sport and Exercise Psychology*, 35(4), 419–432. <https://doi.org/10.1123/jsep.35.4.419>
- Hystad, S. W., Mearns, K. J., & Eid, J. (2014). Moral disengagement as a mechanism between perceptions of organisational injustice and deviant work behaviours. *Safety Science*, 68, 138–145. <https://doi.org/10.1016/j.ssci.2014.03.012>
- Ickes, W., Park, A., & Johnson, A. (2012). Linking identity status to strength of sense of self: Theory and validation. *Self and Identity*, 11(4), 531–544. <https://doi.org/10.1080/15298868.2011.625646>
- Jiang, T., Wang, T., Poon, K. T., Gaer, W., & Wang, X. (2023). Low self-concept clarity inhibits self-control: The mediating effect of global self-continuity. *Personality and Social Psychology Bulletin*, 49(11), 1587–1600. <https://doi.org/10.1177/01461672221109664>
- Khan, U., & Dhar, R. (2006). Licensing effect in consumer choice. *Journal of Marketing Research*, 43(2), 259–266. <https://doi.org/10.1509/jmkr.43.2.259>
- Krol, S. A., & Bartz, J. A. (2022). The self and empathy: Lacking a clear and stable sense of self undermines empathy and helping behavior. *Emotion*, 22(7), 1554–1571. <https://doi.org/10.1037/emo0000943>
- Krol, S. A., Thériault, R., Olson, J. A., Raz, A., & Bartz, J. A. (2020). Self-concept clarity and the bodily self: Malleability across modalities. *Personality and Social Psychology Bulletin*, 46(5), 808–820. <https://doi.org/10.1177/0146167219879126>
- Lamm, C., Bukowski, H., & Silani, G. (2016). From shared to distinct self-other representations in empathy: Evidence from neurotypical function and socio-cognitive disorders. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 371(1686), 20150083. <https://doi.org/10.1098/rstb.2015.0083>

- Lewandowski, G. W., Jr., & Nardone, N. (2012). Self-concept clarity's role in self-other agreement and the accuracy of behavioral prediction. *Self and Identity*, 11(1), 71–89. <https://doi.org/10.1080/15298868.2010.512133>
- Lewandowski, G. W., Jr., Nardone, N., & Raines, A. J. (2010). The role of self-concept clarity in relationship quality. *Self and Identity*, 9(4), 416–433. <https://doi.org/10.1080/15298860903332191>
- Lian, H., Huai, M., Farh, J. L., Huang, J. C., Lee, C., & Chao, M. M. (2022). Leader unethical pro-organizational behavior and employee unethical conduct: Social learning of moral disengagement as a behavioral principle. *Journal of Management*, 48(2), 350–379. <https://doi.org/10.1177/0149206320959699>
- Light, A. E. (2017). Self-concept clarity, self-regulation, and psychological well-being. In J. Lodi-Smith & K. DeMarree (Eds.), *Self-concept clarity: Perspectives on assessment, research, and applications* (pp. 177–193). Springer. https://doi.org/10.1007/978-3-319-71547-6_10
- Light, A. E., & Goldberg, M. H. (2020). Failure to share reality and its consequences for self-concept clarity. *Social Cognition*, 38(1), 62–96. <https://doi.org/10.1521/soco.2020.38.1.62>
- Lodi-Smith, J., & DeMarree, K. G. (2017). Common themes and future directions for self-concept clarity research. *Self-concept clarity: Perspectives on assessment, research, and applications*. Springer. <https://doi.org/10.1007/978-3-319-71547-6>
- Markus, H., & Wurf, E. (1987). The dynamic self-concept: A social psychological perspective. *Annual Review of Psychology*, 38(1), 299–337. <https://doi.org/10.1146/annurev.ps.38.020187.001503>
- McIntyre, K. P., Mattingly, B. A., Lewandowski, G. W., Jr., & Simpson, A. (2014). Workplace self-expansion: Implications for job satisfaction, commitment, self-concept clarity, and self-esteem among the employed and unemployed. *Basic and Applied Social Psychology*, 36(1), 59–69. <https://doi.org/10.1080/01973533.2013.856788>
- Milgram, S. (1963). Behavioral study of obedience. *Journal of Abnormal Psychology*, 67(4), 371–378. <https://doi.org/10.1037/h0040525>
- Mittal, B. (2015). Self-concept clarity: Exploring its role in consumer behavior. *Journal of Economic Psychology*, 46, 98–110. <https://doi.org/10.1016/j.joep.2014.11.003>
- Moore, C. (2015). Moral disengagement. *Current Opinion in Psychology*, 6, 199–204. <https://doi.org/10.1016/j.copsyc.2015.07.018>
- Moore, C., Detert, J. R., Klebe Treviño, L., Baker, V. L., & Mayer, D. M. (2012). Why employees do bad things: Moral disengagement and unethical organizational behavior. *Personnel Psychology*, 65(1), 1–48. <https://doi.org/10.1111/j.1744-6570.2011.01237.x>
- Moose, C. (2008). Moral disengagement in processes of organizational corruption. *Journal of Business Ethics*, 80(1), 129–139. <https://doi.org/10.1007/s10551-007-9447-8>
- Nezlek, J. B., & Plesko, R. M. (2001). Day-to-day relationships among self-concept clarity, self-esteem, daily events, and mood. *Personality and Social Psychology Bulletin*, 27(2), 201–211. <https://doi.org/10.1177/0146167201272006>
- Ogunfowora, B. T., Nguyen, V. Q., Steel, P., & Hwang, C. C. (2022). A meta-analytic investigation of the antecedents, theoretical correlates, and consequences of moral disengagement at work. *Journal of Applied Psychology*, 107(5), 746–775. <https://doi.org/10.1037/apl0000912>
- Organ, D. W. (1973). Locus of control and clarity of self-concept. *Perceptual and Motor Skills*, 37(1), 100–102. <https://doi.org/10.2466/pms.1973.37.1.100>
- Parise, M., Canzi, E., Olivari, M. G., & Ferrari, L. (2019). Self-concept clarity and psychological adjustment in adolescence: The mediating role of emotion regulation. *Personality and Individual Differences*, 138, 363–365. <https://doi.org/10.1016/j.paid.2018.10.023>
- Prusaczyk, E., & Hodson, G. (2020). The roles of political conservatism and binary gender beliefs in predicting prejudices toward gay men and people who are transgender. *Sex Roles*, 82(7–8), 438–446. <https://doi.org/10.1007/s11199-019-01069-1>
- Quinones, C., & Kakabadse, N. K. (2015). Self-concept clarity and compulsive Internet use: The role of preference for virtual interactions and employment status in British and North-American samples. *Journal of Behavioral Addictions*, 4(4), 289–298. <https://doi.org/10.1556/2006.4.2015.038>
- Rahimi, A. M., & Strube, M. J. (2007). Personal self-esteem, collective self-esteem, and self-concept clarity as moderators of the impact of perceived consensus on stereotypes. *Social Influence*, 2(1), 55–79. <https://doi.org/10.1080/15534510701202286>
- Rai, T. S., & Fiske, A. P. (2011). Moral psychology is relationship regulation: Moral motives for unity, hierarchy, equality, and proportionality. *Psychological Review*, 118(1), 57–75. <https://doi.org/10.1037/a0021867>
- Ritchie, T. D., Sedikides, C., Wildschut, T., Arndt, J., & Gidron, Y. (2011). Self-concept clarity mediates the relation between stress and subjective well-being. *Self and Identity*, 10(4), 493–508. <https://doi.org/10.1080/15298868.2010.493066>
- Robins, R. W., Hendin, H. M., & Trzesniewski, K. H. (2001). Measuring global self-esteem: Construct validation of a single-item measure and the Rosenberg Self-Esteem Scale. *Personality and Social Psychology Bulletin*, 27(2), 151–161. <https://doi.org/10.1177/0146167201272002>
- Rozenkrants, B., Wheeler, S. C., & Shiv, B. (2017). Self-expression cues in product rating distributions: When people prefer polarizing products. *Journal of Consumer Research*, 44(4), 759–777. <https://doi.org/10.1093/jcr/ucx067>
- Sammani, A. K., Salamon, S. D., & Singh, P. (2014). Negative affect and counterproductive workplace behavior: The moderating role of moral disengagement and gender. *Journal of Business Ethics*, 119(2), 235–244. <https://doi.org/10.1007/s10551-013-1635-0>
- Sarial-Abi, G., Gürhan-Canli, Z., Kumkale, T., & Yoon, Y. (2016). The effect of self-concept clarity on discretionary spending tendency. *International Journal of Research in Marketing*, 33(3), 612–623. <https://doi.org/10.1016/j.ijresmar.2015.09.010>
- Savary, J., & Dhar, R. (2020). The uncertain self: How self-concept structure affects subscription choice. *The Journal of Consumer Research*, 46(5), 887–903. <https://doi.org/10.1093/jcr/ucz022>
- Savary, J., & Goldsmith, K. (2020). Unobserved altruism: How self-signaling motivations and social benefits shape willingness to donate. *Journal of Experimental Psychology: Applied*, 26(3), 538–550. <https://doi.org/10.1037/xap0000261>
- Servidio, R., Sinatra, M., Griffiths, M. D., & Monacis, L. (2021). Social comparison orientation and fear of missing out as mediators between self-concept clarity and problematic smartphone use. *Addictive Behaviors*, 122, Article 107014. <https://doi.org/10.1016/j.addbeh.2021.107014>
- 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. <https://doi.org/10.1177/0146167211398138>
- Smith, R. H., Webster, J. M., Parrott, W. G., & Eyre, H. L. (2002). The role of public exposure in moral and nonmoral shame and guilt. *Journal of Personality and Social Psychology*, 83(1), 138–159. <https://doi.org/10.1037/0022-3514.83.1.138>
- Steele, C. M. (1988). The psychology of self-affirmation: Sustaining the integrity of the self. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 21, pp. 261–302). Academic Press. [https://doi.org/10.1016/S0065-2601\(08\)60229-4](https://doi.org/10.1016/S0065-2601(08)60229-4)
- Strohming, N., & Nichols, S. (2014). The essential moral self. *Cognition*, 131(1), 159–171. <https://doi.org/10.1016/j.cognition.2013.12.005>
- Teng, Z., Nie, Q., Zhu, Z., & Guo, C. (2020). Violent video game exposure and (Cyber) bullying perpetration among Chinese youth: The moderating role of trait aggression and moral identity. *Computers in Human Behavior*, 104, Article 106193. <https://doi.org/10.1016/j.chb.2019.106193>
- Tesser, A. (1988). Toward a self-evaluation maintenance model of social behavior. In L. Berkowitz (Ed.), *Advances in experimental social psychology*

- (Vol. 21, pp. 181–227). Academic Press. [https://doi.org/10.1016/S0065-2601\(08\)60227-0](https://doi.org/10.1016/S0065-2601(08)60227-0)
- Thornberg, R., & Jungert, T. (2014). School bullying and the mechanisms of moral disengagement. *Aggressive Behavior*, 40(2), 99–108. <https://doi.org/10.1002/ab.21509>
- Touré-Tillery, M., & Fishbach, A. (2012). The end justifies the means, but only in the middle. *Journal of Experimental Psychology: General*, 141(3), 570–583. <https://doi.org/10.1037/a0025928>
- Touré-Tillery, M., & Fishbach, A. (2015). It was(n't) me: Exercising restraint when choices appear self-diagnostic. *Journal of Personality and Social Psychology*, 109(6), 1117–1131. <https://doi.org/10.1037/a0039536>
- Touré-Tillery, M., & Light, A. E. (2018). No self to spare: How the cognitive structure of the self influences moral behavior. *Organizational Behavior and Human Decision Processes*, 147, 48–64. <https://doi.org/10.1016/j.obhdp.2018.05.002>
- Touré-Tillery, M., & Wang, L. (2022). The good-on-paper effect: How the decision context influences virtuous behavior. *Marketing Science*, 41(5), 1004–1024. <https://doi.org/10.1287/mksc.2021.1347>
- Vincent, L. C., Emich, K. J., & Goncalo, J. A. (2013). Stretching the moral gray zone: Positive affect, moral disengagement, and dishonesty. *Psychological Science*, 24(4), 595–599. <https://doi.org/10.1177/0956797612458806>
- Wang, J., & Yu, Y. (2023). Beautify the blurry self: Low self-concept clarity increases appearance management. *Journal of Consumer Psychology*, 33(2), 377–393. <https://doi.org/10.1002/jcpy.1298>
- Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology*, 54(6), 1063–1070. <https://doi.org/10.1037/0022-3514.54.6.1063>
- Wolf, N. (2013). *The beauty myth: How images of beauty are used against women*. Random House.
- Zheng, X., Qin, X., Liu, X., & Liao, H. (2019). Will creative employees always make trouble? Investigating the roles of moral identity and moral disengagement. *Journal of Business Ethics*, 157(3), 653–672. <https://doi.org/10.1007/s10551-017-3683-3>

Received July 3, 2023

Revision received May 15, 2024

Accepted July 5, 2024 ■