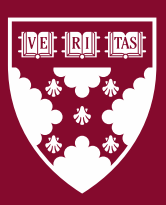
Working Paper 23-045

"It Wouldn't Have Mattered Anyway": When Overdetermined Outcomes Justify Our Sins

Stephanie C. Lin
Julian J. Zlatev
Dale T. Miller



Harvard Business School

“It Wouldn’t Have Mattered Anyway”: When Overdetermined Outcomes Justify Our Sins

Stephanie C. Lin
INSEAD

Julian J. Zlatev
Harvard Business School

Dale T. Miller
Stanford University

Working Paper 23-045

Funding for this research was provided in part by Harvard Business School.

"It wouldn't have mattered anyway": When overdetermined outcomes justify our sins

Stephanie C. Lin

INSEAD

Julian J. Zlatev

Harvard University

Dale T. Miller

Stanford University

Abstract

We identify and document an "overdetermined outcome defense" which occurs when one learns that circumstances besides one's own actions were sufficient to produce a negative effect (e.g., deciding not to go to the gym, but later discovering that the gym had been closed anyway). We present seven preregistered studies (total $N = 3784$) examining the nature of this effect. In Studies 1 and 2 we find people felt less guilty when they discovered that a negative outcome following a self-standard violation "would have happened anyway" than if no such information was available across a variety of scenarios (Study 1) and spontaneously in a thought generation paradigm (Study 2). Studies 3 and 4 suggest this effect has both rational and motivated components. In Study 3, the overdetermined outcome defense was used for both others and the self, but more for the self, and in Study 4, the participants admitted that they would use the overdetermined outcome defense more than they felt they should. In Studies 5A, 5B and 6, we tested behavioral implications of this effect in self-control and prosocial behavior contexts. We found that, when people chose options that violate their self-standards over ones more consistent with their self-standards, guilt motivated them to actively seek overdetermination to justify their choice. In Study 6, we once again found that discovering that a negative outcome was overdetermined reduced guilt in participants in an incentive-compatible decision context. Our findings contribute to the literatures on outcome bias, justification, and irrational acquiescence.

Keywords: guilt avoidance and justification; morality and prosocial behavior; self-control; should/want conflict; decision making

Imagine that you have been planning on heading to the gym after work. However, when the time comes you decide to instead go home and binge watch TV. While doing so, it occurs to you that the gym may actually be closed due to renovations. Are you tempted to check whether the gym is closed? If it is closed, are you less culpable than you originally thought for your sinful behavior? In this paper, we identify and explore the psychological boundaries of a phenomenon that we call the overdetermined outcome defense. We suggest that when the consequences of one's sins are overdetermined—that is, when one's actions were not necessary to result in the negative consequences that followed—people feel less guilty about their sins. This leads people to seek information of this type, even while acknowledging that the comfort it provides is illegitimate.

Below, we define overdetermined outcomes and the overdetermined outcome defense. Next, we argue for a hybrid motivated-cognitive process for the overdetermined outcome defense. We then outline predictions on how and when the overdetermined outcome defense should lead to information seeking. Finally, we test our outlined propositions in seven empirical studies.

Defining Overdetermined Outcomes

We define an overdetermined outcome as an outcome that a person's actions were *intended* and *sufficient* to cause, but that would have occurred even without those actions. In other words, an outcome is overdetermined when there exists at least one other sufficient cause beyond one's own actions. In the introductory example, the person's decision to stay home rather than go to the gym is *sufficient* to cause them to not exercise that day. However, in this case not exercising is also an *overdetermined* outcome; even if they had decided to go to the gym, at least one other factor—namely, the gym being closed—was also sufficient to cause the non-exercising

outcome. We define the "overdetermined outcome defense" as the tendency for people to use the overdetermination of the negative outcome that their intended actions were sufficient to cause in order to decrease the perceived culpability of their purposefully sinful action (e.g., in this case, one's decision to not go to the gym). In the cases we consider, the person's intention to produce the negative outcome is unambiguous; the intentionality of the act cannot be informed by the discovery of the overdetermining factor[1]. Returning to our opening example, the person is fully aware that they intended not to exercise, and that this outcome would have occurred whether or not the gym was closed. Regardless of this fact, we suggest that the person would feel better about not going to the gym after finding out the gym was closed. We outline motivated and cognitive processes contributing to this defense below.

The Overdetermined Outcome Defense as a Motivated-Cognitive Process

Motivated Components. We suggest that people are especially motivated to use the overdetermined outcome defense, specifically to justify their own sinful behavior in the context of want/should conflicts. People often face conflicts between what they *want* to do and what they feel they *should* do in both the self-control and moral domains (Bitterly, Mislavsky, Dai, & Milkman, 2015; Milkman, Rogers, & Bazerman, 2008). In the self-control domain, people may *want* to eat decadent chocolate cake and binge watch TV but, to adhere to their long-term health goals, instead feel that they *should* order fruit salad or exercise (Khan, Dhar, & Wertenbroch,

[1] This stands in contrast to two well-known phenomena: (1) the discounting principle of attribution theory, which states that when more than one possible cause of an effect are identified, the causal role of the any given cause is discounted (Kelley, 1973); related work further argues that people are less likely to attribute outcomes to internal sources when evidence for external sources becomes more salient (Kruglanski, 1980; Weiner, 1985) and (2) outcome bias, in which observers view actions that produce a negative outcome more negatively than identical actions that do not (e.g., Agrawal & Maheswaran, 2005; Baron & Hershey, 1988; Gino et al., 2010; Shaver, 1970). Critically, attribution theory focuses on instances in which there is initial ambiguity around the potential cause of a behavior. In the same vein, outcome bias can be attributed to an overapplication of a general rule that outcomes inform the quality of the decision (Baron & Hershey, 1988). In the case of overdetermined outcomes, the outcome clearly does not inform the decision. Furthermore, neither the discounting principle nor outcome bias are considered to be motivated. As we will argue, the overdetermined outcome defense is at least partially motivated.

2005; Wertenbroch, 1998). Similarly, in the domain of moral behavior, people must often choose between engaging in self-benefitting behaviors they want to do and other-benefitting behaviors they feel they should do to benefit society (e.g., protecting the environment, donating to charitable organizations, or people in need) (Berman & Small, 2012; Liu & Lin, 2018; Reczek, Irwin, Zane, & Ehrich, 2017; Zlatev, Kupor, Laurin, & Miller, 2020). For instance, people may want to buy less expensive but environmentally unsustainable shampoo yet feel that they should buy the more expensive and sustainable alternative; or they may want to stay at home and watch TV but feel that they should instead volunteer at the local homeless shelter. In this article, we refer to the decision to engage in *want* behaviors (over *should* behaviors) in either the self-control or moral domain as self-standard violations and characterize such violations and their consequences as negative or sinful.

When people engage in these self-standard violations, they feel guilty[2], a state that people generally are motivated to avoid (Baumeister, Stillwell, & Heatherton, 1994; Dahl, Honea, & Manchanda, 2003; Tangney, Stuewig, & Mashek, 2007; Wertenbroch, 1998; Xu & Schwarz, 2009). One way people ward off guilt following self-standard violations is by justifying them (Huberts, Evers, & Ridder, 2014; Khan & Dhar, 2006; Kivetz & Zheng, 2006; Lin & Miller, 2021; Lin, Zlatev, & Miller, 2017; Liu & Lin, 2018; Okada, 2005; Shalvi, Gino, Barkan, & Ayal, 2015). The justifications people resort to for actions that have the *prima facie* appearance of violations show people to be casuists: judges who assess the virtuousness of an action by focusing on its particular circumstances rather than simply applying a general rule or principle to it (Norton, Vandello, & Darley, 2004; Uhlmann & Cohen, 2005). Case-specific principles that

[2] While guilt is only one negative reaction people may have to self-standard violations, it is a typical affective response indicating negative self-evaluation that occurs after both moral and self-control failures (Baumeister & Heatherton, 1996). As a result, we focus on guilt in this paper, though acknowledging similar relationships may exist with other emotions, such as shame and regret.

people evoke to reduce guilt from sinful *want* behaviors include the following: one can hurt others if they deserve it (Bandura, Barbaranelli, Caprara, & Pastorelli, 1996); one cannot be blamed for actions they had little choice in taking (Malle & Knobe, 1997); and one is entitled to cheat if one shares the spoils (Wiltermuth, 2011). People may also exaggerate the relative virtuousness of their past behavior (Effron, Miller, & Monin, 2012; May & Irmak, 2014) or promise themselves that they will make up for their bad behavior in the future (Cascio & Plant, 2014; Khan & Dhar, 2007).

These justification strategies are examples of classic self-serving attributions whereby people attribute their own failures to external circumstances (Kunda, 1990; Lin et al., 2017; Meyer, 1980; Zuckerman, 1979) to feel better about failures (McFarland & Ross, 1982). In all these cases, people are motivated to legitimate the exclusion of self-interested or indulgent behavior from self-evaluation on the basis of mitigating circumstances that they feel render their self-standard violations morally uninformative. Rendering their self-standard violations situationally permissible is an important means, but not the only one, that people have to reduce their guilt. We propose that people can also avoid feeling bad about the negative consequences of their self-standard violations—even those they admit were ill-intended—if they can claim that the outcome of their violation behavior was overdetermined. Even knowing that one's own decision was sufficient in causing the negative outcome, an additional circumstance that would have led to the same outcome (i.e., an overdetermining factor) discovered retroactively can mitigate negative self-judgment. This leads to our first hypothesis:

H1: People feel lower negative self-evaluation about their sinful decisions when the outcomes of those decisions are overdetermined than when they are not.

A hybrid motivated-cognitive process. Although the present research focuses on the motivated use of the overdetermined outcome defense, there may be cognitive components to the

process as well. Specifically, the reasoning behind the overdetermined outcome defense is a form of counterfactual thinking—that is, imagining what would have happened if one had *not* made the decision that one did (Miller, Turnbull, & McFarland, 1990; Roese, 1997). Causal inferences are often made when one mentally alters a situation (for instance, deleting or altering an action) and simulates the outcome (Kahneman & Miller, 1986); if the mental alteration of the situation changes the outcome, one can infer that the altered action caused the original outcome (Pearl, 2009). For instance, absent overdetermination, someone who decided to stay home from the gym might imagine what would have happened if she had made a different decision. In that case, she would have gone to the gym; thus, her decision to stay home caused her not to exercise (Roese, 1997). In the case of an overdetermined outcome, however, mentally simulating this scenario does not lead to an *undoing* of the consequence. If the gym was closed, then even changing the decision to stay home would not have led to a session at the gym. That is, the answer to the question "who is to blame that one did not work out today" has a more ambiguous answer in the overdetermined case than in the non-overdetermined case. Thus, people can make a relatively more compelling argument that they are not fully to blame for *the consequence* of their decision when the outcome is overdetermined than when it is not.

Taking the motivated and cognitive components of the overdetermined outcome defense together, we suggest that people endorse the use of the overdetermined outcome defense, reflecting a cognitive process, but extend their use of the defense beyond their own endorsement, reflecting a motivated process. That is, people view the palliative use of overdetermined outcomes as semi-legitimate.

Specifically, if use of the overdetermined outcome defense were a purely cognitive process, we would expect that it would alleviate judgment of others' violation behaviors as much

as it alleviated one's own. On the other hand, if the overdetermined outcome defense is a motivated extension of a cognitive process, we would expect an actor-observer discrepancy in the use of overdetermined outcomes: overdetermined outcomes should alleviate the severity of judgments of others' sins, reflecting a cognitive process, but should alleviate the severity of one's own sins more, reflecting an additional motivated process. Formally:

H2: People display an actor-observer bias such that overdetermined outcomes excuse others' behavior but excuse their own behavior to a greater extent.

Furthermore, we suggest that people are aware that their use of the overdetermined outcome bias is "consolation overreach": Specifically, although people believe that an overdetermined outcome should somewhat alleviate their guilt (reflecting a cognitive process), it *actually* alleviates guilt to a greater extent than they believe it *should* (reflecting a motivated process). Furthermore, we suggest that people consciously employ the overdetermined outcome defense and do so despite recognizing its questionable legitimacy. In this respect, the phenomenon comports with recent work on acquiescence, wherein people identify, but fail to correct for, errors in their judgment (Risen, 2015; Walco & Risen, 2017). Specifically:

H3: People believe that overdetermined outcomes should alleviate their guilt to a certain extent but not to the extent that it does alleviate their guilt.

Information Seeking as a Behavioral Consequence

Because overdetermined outcomes alleviate the perceived negativity of one's self-standard violations, people may be motivated to actively seek evidence that an outcome that followed a sinful act was overdetermined. This type of information seeking stands in contrast to classic information avoidance (Golman, Hagmann, & Loewenstein, 2017) in which people avoid potentially useful information to protect themselves. We propose that people may spend resources (e.g., time) to seek normatively useless information to reduce guilt. For example,

finding out that the gym was closed has no impact on one's prior decision not to go to the gym, and makes no consequential difference on the outcome of one's decision; thus it is wasteful to spend one's time and energy seeking such information. Just as people engage in biased information seeking that confirms their decisions (Jonas, Schulz-Hardt, Frey, & Thelen, 2001; Schulz-Hardt, Frey, Lüthgens, & Moscovici, 2000), we suggest that people will expend effort to find out whether or not the negative consequences of their *want* behaviors actually materialized (more so than for their *should* behaviors). Formally:

H4: When people have engaged in sinful (vs. virtuous) behavior, they seek information that could reveal that an outcome is overdetermined in order to alleviate their guilt.

A strict interpretation of the information avoidance literature suggests that, when faced with the possibility that the outcome of their sinful decision may have been overdetermined, people would avoid further information, as finding out that the consequence was not overdetermined would provide evidence of the consequential nature of their decision. However, we argue that in situations when one has already clearly made a sinful decision, people have more to gain by finding out that their decision was overdetermined than they have to lose by finding out that their decision was not overdetermined. In other words, we suggest that the baseline assumption is already that one's decision had a causal impact; confirming that it indeed had causal impact may make people feel somewhat worse but finding out that it did *not* have causal impact would make people feel much better. In a sense, people's self-evaluation can go further up than it can go down. Thus:

H5: People adhere to H4 at both low and high probabilities that one's decision is overdetermined.

The Present Research

In summary, we propose that knowing that one's violation of a self-standard was not

necessary to produce a projected negative outcome alleviates guilt (i.e., "the overdetermined outcome defense"). We further propose that this process involves both cognitive and motivated components. We present three sets of preregistered studies that document the nature of this effect (see Table 1). The first set of studies tests H1: Study 1 tests whether, across a variety of scenarios, discovering that one's self-standard violation was not pivotal in producing negative consequences alleviates guilt; Study 2 tests whether this alleviation of guilt occurs spontaneously in participant's self-generated thoughts.

The second set of studies examine the cognitive and motivated processes involved in this effect. Specifically, Study 3 tests whether overdetermined outcomes alleviate judgments of others' sinful decisions, reflecting a cognitive process, but alleviate judgments of their own sinful decisions more, reflecting a motivated process (H2). Study 4 tests whether following sinful behavior, people believe that an overdetermined outcome should somewhat alleviate their guilt (reflecting a cognitive process) but that it alleviates guilt more than they believe it should (reflecting a motivated process; H3).

The final set of studies examines active information seeking behavior regarding overdetermination of their outcomes. In Studies 5A and 5B, we examine whether people pay a real time cost to find out whether their sinful (vs. virtuous) choice was overdetermined. We predict an increase in information seeking after the sinful (vs. virtuous) decision (H4). Study 5B manipulated the probability that the outcome was overdetermined, testing whether people are insensitive to such information when seeking outcomes (H5). Finally, Study 6 tests the overdetermined outcome defense in a behavioral setting in which participants have chosen to engage in self-interested (over prosocial) behavior.

Table 1

Summary of studies

	Study	IV	DV	Main Finding and Hypotheses Tested
Studies examining overdetermined outcome defense	1	Outcome (overdetermined, control) × Scenario (recycling, green behavior, volunteering, healthy eating, exercising)	Guilt	Across five scenarios in both self-control and moral domains, people felt less guilty when they find out that a sinful decision was overdetermined than when it was not overdetermined (H1).
	2	Outcome (overdetermined, control)	Self-classification of spontaneous thoughts (i.e., as "feeling better")	After finding out a sinful decision was overdetermined (vs. not) participants were more likely to list thoughts that they self-classify as "feeling better" (H1); a high ratio of participants in the overdetermined condition indicated feeling better (78%).
Studies addressing cognitive and motivated processes	3	Outcome (overdetermined, control) × target (self, other)	Judgment of virtuousness	Overdetermined outcomes (vs. control) increased judgments of virtue for both others and the self, but the effect was larger for the self (H2).
	4	Outcome (overdetermined, control) × guilt type (should, actual)	Guilt	People thought they should feel somewhat less guilty when an outcome was overdetermined (vs. control); however, they thought that they should not feel less guilty to the extent that they actually would feel less guilty (H3).
Effects on information seeking	5A	Choice (violation, non-violation)	Seeking overdetermined outcome information	Those who made a sinful decision (vs. those who made a virtuous decision) were more likely to seek overdetermined outcome information, which was mediated by guilt (H4).
	5B	Choice (violation, non-violation) × overdetermination chance (10, 20, 30, 40, 50, 60, 70, 80, 90%)	Seeking overdetermined outcome information	Those who made a sinful decision (vs. those who made a virtuous decision) were more likely to seek overdetermined outcome information (H4). True at both and high chance of overdetermination (H5); suggestive evidence that the effect may be weaker at moderate chance of overdetermination (i.e., highest uncertainty).
	6	1. Outcome (overdetermined, non-overdetermined) × Guilt (before, after outcome reveal), mixed; 2. Measured guilt about sinful decision	1. Guilt, 2. Information seeking	1. Those who learned that their outcome was overdetermined felt a decrease in guilt (H1) whereas those who learned that their outcome was not overdetermined felt an increase in guilt. 2. Measured guilt about a sinful decision led people to seek information about whether their outcome was overdetermined (H4).

All studies were preregistered and included at least 100 participants per condition (Simmons, Nelson, & Simonsohn, 2018). Data were collected from Prolific (Studies 1–5) and Amazon's Mechanical Turk (Study 6). Online panels such as these have been shown to yield diverse samples in age, education level, and economic strata in a US cultural context (Paolacci & Chandler, 2014). We report all measures and conditions included in each study. Materials, data, R scripts, and preregistrations for all studies can be viewed here: https://researchbox.org/287&PEER_REVIEW_passcode=BMBGMQ. All studies received

Institutional Review Board approval.

Study 1

In Study 1, we tested H1, which states that people feel less negative self-evaluation (here, guilt) about self-standard violations when the outcomes that follow are overdetermined. Specifically, people imagined engaging in a variety of self-standard violations (i.e., immoral, self-interested, or indulgent decisions). They were then randomly assigned to find out that the consequences of those violations were or were not overdetermined—that is, whether these outcomes "would have happened anyway." We predicted that those who discovered that the negative consequences of their self-standard violations were overdetermined would feel less guilty than those who did not.

Method

Participants and Design

We preregistered this study and opened it to 1,000 participants on Prolific, which resulted in 1,001 participants (M_{age} = 31.70; 528 male, 472 female, 1 did not respond). This study was a 2 (outcome: overdetermined vs. non-overdetermined control) × 5 (scenario: recycling, green behavior, volunteering, healthy eating, exercising) between-subjects design.

Procedure

Participants were randomly assigned to one of five scenarios in which they imagined that they chose (1) not to recycle, (2) not to purchase an eco-friendly shampoo, (3) not to volunteer at a homeless shelter, (4) not to exercise, or (5) to order a cookie rather than a fruit salad (see Table 2 for full text). They were then randomly assigned to discover that the negative outcome of their choice was consequential (control condition) or that the negative outcome was overdetermined and would have happened anyway (overdetermined condition). Specifically, they imagined,

respectively, that (overdetermined condition in brackets) (1) the recycling was taken to the recycling plant [landfill], (2) the eco-friendly shampoo had been in stock [out of stock], (3) the homeless shelter was open [closed] during the hours they would have volunteered, (4) the gym was open [closed] during the hours they would have gone to the gym, and (5) the fruit salad was in stock [out of stock].

Table 2

Scenarios presented in Study 1

Scenario	Scenario text	Manipulation [overdetermined condition in brackets]
Recycling	Imagine that you just finished drinking a cold soda on a hot day at the park. You look around and see a trash can a few feet away from you. You see that the nearest recycling would require you to walk all the way across the park. You decide to throw your soda can out in the trash bin.	You find out that the recycling at this park is actually taken to the recycling plant [to the landfill rather than to the recycling plant].
Green behavior	Imagine that you are looking for a new shampoo online, and are deciding between an eco-friendly brand that is slightly more expensive and a non-eco-friendly brand that is cheaper. You choose the non-eco-friendly shampoo. The online store sometimes runs out of shampoo, and therefore you may not have been able to order the eco-friendly shampoo anyway. You did not know whether the eco-friendly shampoo was in-stock or out of stock when you chose to order the non-eco-friendly shampoo.	After you finish ordering, you find out that the eco-friendly shampoo had been in stock [out of stock], meaning you could have [could not have] ordered it.
Volunteering	Imagine that, a few days ago, you saw a flyer for the homeless shelter in your area, which is looking for volunteers for their food service. You are deciding whether to go volunteer this afternoon. You decide that you don't feel like going. Later that night, you realize that you may have misremembered the hours for the shelter and that it might not have even been open during the time you had been thinking about going.	You check online and find out that the shelter was actually open [closed] during the time you would have gone.
Exercising	Imagine that you were deciding whether to go to the gym this afternoon. You decide that you don't feel like going.	You check online and find out that the gym was actually open [closed] during the time you would have gone.

	Later that night, you realize that you may have misremembered the hours for the gym and that it might not have even been open during the time you had been thinking about going.	
Healthy eating	Imagine that you have recently been trying to eat healthy. You are in line at a cafe and want to order a snack. The cafe has a cookie and a fruit salad on the menu, and you are choosing between the two. You decide to order the cookie. The cafe sometimes runs out of snacks, and therefore you may not have been able to order the fruit salad anyway. You did not know whether the fruit salad was in-stock or out of stock when you chose to order the cookie.	After you finish eating, you find out that the fruit salad had been in stock [out of stock], meaning you could [could not] have ordered it.

Participants then indicated how guilty they would feel about their decision in this situation (1 = *not at all*, 5 = *extremely*).

Results

We regressed ratings of guilt on outcome condition. Results revealed a significant effect of guilt, such that participants in the overdetermined condition ($M = 2.18$, $SD = 1.50$) indicated they would feel significantly less guilty than participants in the control condition ($M = 3.35$, $SD = 1.50$), $t(999) = -12.37$, $p < .001$, d = .78.

As preregistered, we next regressed guilt onto outcome condition, scenario, and their interaction. As indicated in the preregistration, we were not interested in a condition by scenario interaction (as it would not be surprising if the effect were smaller or larger in any given scenario). However, this analysis allowed us to include scenario fixed effects, to control for any differences in by-scenario intercepts, and to examine the effect of condition within each scenario by dummy coding each scenario as the reference scenario. This revealed that the difference by overdetermined outcome condition was significant for each individual scenario, all *p*s < .001 (see Table 2).

Finally, as an exploratory analysis, we examined whether there were any interactions between overdetermined outcome condition and scenario type. The overall 2 (outcome condition) × 5 (scenario) interaction was not significant, $F(4, 991) = 1.24$, $p = .293$. Dummy coding each scenario as the reference condition to compare effects of the outcome conditions between scenarios revealed only one significant difference: the effect of outcome condition in the recycling scenario was larger than the effect of outcome condition in the healthy eating scenario, B = .58, $t(991) = 1.96$, $p = .050$.

Table 2

Descriptive and Test Statistics of Condition on Guilt in Study 1 for Each Scenario

Scenario	Overdetermined Condition	Control Condition	*t*-values (df = 991)	*p*-value	Cohen's *d*
Recycling	2.68 (1.69)	4.13 (1.43)	–7.01	<.001	.92
Green behavior	1.93 (1.27)	3.09 (1.56)	–5.65	<.001	.82
Volunteering	2.39 (1.59)	3.36 (1.21)	–4.71	<.001	.68
Exercising	1.72 (1.24)	2.98 (1.48)	–6.13	<.001	.93
Healthy eating	2.19 (1.54)	3.07 (1.49)	–4.22	<.001	.58

Discussion

This study provides evidence that people anticipate that overdetermined outcomes will mitigate their guilt for engaging in behaviors that violate their self-standards relative to when those outcomes were not overdetermined (H1). That is, when people discovered that the consequences that followed their self-standard violating behaviors would have happened irrespective of those behaviors, they felt less guilty. We demonstrated this effect across a variety of scenarios in both the moral and self-control domains. However, this study relied on elicitation of guilt as evidence of the overdetermined outcome defense. Study 2 tests whether people

spontaneously use this defense when faced with overdetermined outcomes using a thought listing task similar to those used in previous research (Johnson, Häubl, & Keinan, 2007; Zlatev & Miller, 2016).

Study 2

Study 2 tested whether overdetermined outcomes decrease people's guilt about their sinful decisions in their spontaneous reactions rather than on a scale measure of guilt, further testing H1. Participants were faced with one of the scenarios from Study 1 (i.e., the exercising scenario) and again were assigned to either an overdetermined outcome (i.e., the gym was closed) or non-overdetermined outcome control (i.e., the gym was open) condition. Participants then listed their spontaneous reactions to this outcome, and self-classified the thoughts they listed. We hypothesized that participants would indicate that their thoughts reflected "feeling better" in the overdetermined (vs. non-overdetermined) outcome condition.

Method

Participants and Design

We preregistered this study and opened it to 300 participants on Prolific, which resulted in 301 participants (M_{age} = 36.81; 150 male, 143 female, 8 non-binary). This study was a two-condition (overdetermined outcome vs. non-overdetermined outcome control) between-subjects design.

Procedure

Participants were presented with the exercising scenario from Study 1 and were assigned to either the overdetermined or control condition. After reading the assigned scenario, they indicated their thoughts on the situation, and were asked to write at least 50 characters before moving on. On the next page, they were shown the thoughts they had listed and categorized them

into one of three categories: (1) The information I learned made me feel better, (2) The information I learned made me feel worse, or (3) Neither[3].

Results

A Pearson's Chi-square test revealed that the distributions of self-categorizations differed between conditions $X^2(2, N = 301) = 144.32$, $p < .001$. In the overdetermined outcome condition, after learning that the gym was closed, 78.6% of participants classified themselves as feeling better, 5.7% indicated feeling worse, and 15.7% did not classify their thoughts in either category. In contrast, in the control condition, after learning that the gym was open 10.6% classified themselves as feeling better, 46.5% indicated feeling worse, and 44.0% did not classify their thoughts in either category. A binomial logistic regression, regressing self-classification of the "feel better" item (coded as 1) versus "feel worse" and "neither" items (coded as 0) revealed that those in the overdetermined outcome condition were more likely to spontaneously feel better than those in the control condition, B = 3.44, SE = .33, $p < .001$, odds ratio (OR) = 31.23. On the other hand, they were less likely to indicate feeling worse than those in the control condition, B = –2.67, SE = .38, $p < .001$, OR = .07.

Although not a preregistered analysis, we additionally found that a higher ratio of participants indicated feeling better in the overdetermined outcome condition (78.6%) than the ratio of those feeling worse in the control condition (46.7%), $z = 5.78$, $p < .001$.

Discussion

Study 2 provided evidence that people spontaneously feel better after learning that the outcomes of their sinful decisions were overdetermined. Without being primed with the concept

[3] The preregistration specifies a fourth option that we later removed from the study but neglected to remove from the preregistration. It does not affect the data analysis, as we indicated we would analyze binary variables of whether they selected option 1 or not, and whether they selected option 2 or not.

of guilt, participants who learned that their outcome was overdetermined (i.e., the gym had been closed when they would have exercised) organically listed thoughts that they then self-classified as "feeling better" the majority (78.6%) of the time. This was significantly higher than those who learned that their outcome was not overdetermined (i.e., the gym had been open when they would have exercised), who only classified themselves as feeling better 10.6% of the time.

Perhaps more informative, participants whose outcomes were overdetermined felt better more often than those whose outcomes were not overdetermined felt worse (46.7%). This speaks to the fact that those who already made a sinful decision have more room to increase than to decrease their self-evaluation. That is, their baseline assumption was already that they sinned; finding out that the gym had been open (even after briefly thinking that the gym may have been closed) thus did not affect their feelings as much as those who found out the gym had been closed. This supports assumptions related to H5, which we directly test in Study 5B.

Thus Studies 1 and 2 provide evidence for our main hypothesis that overdetermined outcomes can decrease negative self-evaluations about one's sins. Our second set of studies (Studies 3 and 4) examine the cognitive and motivated components of this effect. Specifically, Study 3 tests whether people will apply the overdetermined outcome defense to others to the same extent that they apply it to themselves.

Study 3

Our first two studies demonstrated that people feel less guilty when they discover that the negative outcomes of their self-standard violations are overdetermined (H1). In Study 3, we test H2, which predicts an actor-observer discrepancy (Pronin, Gilovich, & Ross, 2004), such that information regarding whether sinful actions were overdetermined would forgive one's own *want* behavior more than it would forgive the behavior of others. This predicted result would

support the claim that use of such motivated justifications is viewed as semi-legitimate and is resorted to when motivated to protect oneself against self-threat.

Method

Participants and Design

We preregistered this study and opened it to 800 participants from Prolific, resulting in 801 participants (M_{age} = 36.15; 397 male, 385 female, 17 non-binary, 2 did not respond). This study had a 2 (outcome: overdetermined vs. non-overdetermined control) × 2 (target: self vs. other) between-subjects design.

Procedure

Participants read the recycling scenario from Study 1 but imagined themselves (self condition) or another person (John; other condition) in the scenario. As in Study 1, participants then read that they [John] later discovered that the recycling was actually taken to the landfill rather than the recycling plant (overdetermined outcome condition) or taken to the recycling plant (control condition). Participants then indicated the extent to which their [John's] decision to throw the soda out in the trash bin was: bad or good, sinful or virtuous, and wrong or right on 7-point scales (e.g., 1 = *extremely bad*, 7 = *extremely good*), which were averaged as our measure of virtuousness.

Results

We regressed virtuousness onto outcome condition, target condition, and their interaction. People judged John's behavior as more virtuous when his decision was overdetermined (M = 4.61, SD = .99) than when it was not (M = 4.13, SD = 1.17), t(797) = 4.08, $p < .001$, d = .44. People also judged their own behavior as more virtuous when their decision was overdetermined (M = 4.72, SD = 1.10) than when it was not (M = 3.75, SD = 1.37), t(797) =

8.36, $p < .001$, d = .78. However, critically, a significant interaction revealed that this effect was larger for the self-evaluation than for other-evaluation, $t(797) = 3.00$, $p = .003$.[4]

Discussion

In Study 3, we found that the effect of having an overdetermined outcome on judgment was larger for self-evaluations than for other-evaluations (H2). Specifically, although people believed that another person's decision to not recycle a soda can was more virtuous when it would have been taken to the landfill than when it would have been properly recycled, this effect was twice as large for oneself ($B_{other} = .48$ vs. $B_{self} = .97$). This study provides evidence that, although there is an arguably legitimate component of the application of the overdetermined outcome defense, people apply it more when it would relieve their own sense of guilt than in judgment of others' behavior.

In Study 4 we further examine the nature of this bias. To what extent is this behavior self-deceptive? Do people believe their use of the overdetermined defense to justify their own sins is legitimate, or are they aware of their bias?

Study 4

Study 3 demonstrated a motivated bias in the use of overdetermined outcomes in judgments of virtue. In Study 4, we explored whether people themselves view the use of overdetermined outcomes as a legitimate reaction, and if so, to what extent. Considerable previous work has found that people are typically unaware of their own biases (Pronin et al., 2004; Pronin, Lin, & Ross, 2002; Scopelliti et al., 2015). However, research shows that people

[4] Decomposing the interaction the other way, those who judged their own behavior judged the act of not recycling as less virtuous than those who judged others' behavior when there was no overdetermined outcome present, $t(797) = -3.24$, $p = .002$, d = .30. However, when that outcome was overdetermined, there was no such difference, $t(797) = 1.00$, $p = .32$, d = .11. This pattern may be due to the fact that people at baseline believe they are less unethical than others (Klein & Epley, 2016), and therefore a sinful act like not recycling violates expectations more for the self than for others. However, the overdetermined outcome overcame this harsher negative self-judgment.

sometimes will engage in biased behavior or thinking even when they recognize and admit that it is biased (Risen, 2015).

In Study 4, we investigated whether people view their use of an overdetermined outcome justification as consolation overreach. That is, we predicted that they would express the belief that the overdetermined nature of their intended outcome should alleviate guilt but not as much as it actually would (H3).

Method

Participants and Design

We preregistered this study and opened it to 300 participants from Prolific, resulting in 301 participants (M_{age} = 33.35; 159 male, 142 female). This study had a 2 (outcome: overdetermined vs. non-overdetermined control) × 2 (guilt type: should vs. actual) mixed design in which the outcome manipulation was administered between subjects and the guilt type was measured within subject.

Procedure

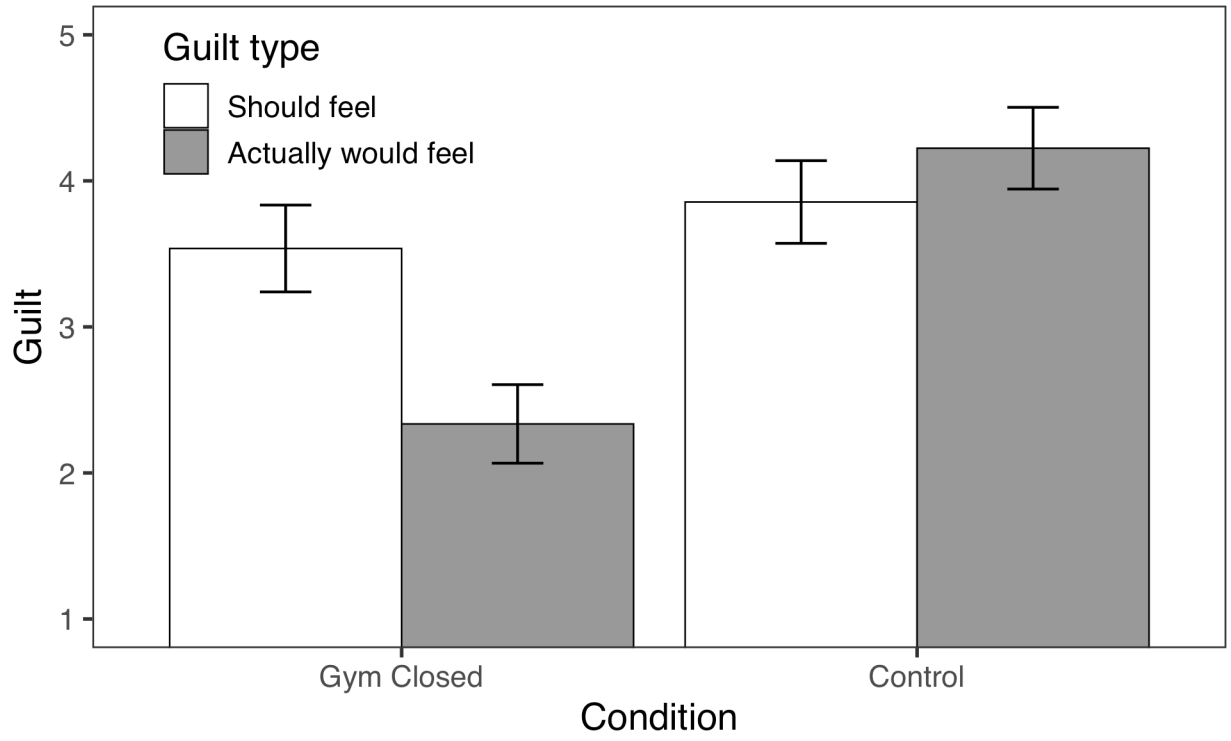
Participants read that they would be rating how they would actually feel and how they should feel in certain situations. They were informed that sometimes the way people actually feel and the way they should feel are different, and that sometimes they are the same, and were asked to answer as accurately and honestly as possible (DeMarree, Wheeler, Briñol, & Petty, 2014). Using a scenario adapted from Study 2, participants then read that they were considering going to the gym but decided to stay home. In the overdetermined outcome condition, they were told that they discovered later that the gym was closed during the time they would have gone anyway; those in the control condition were not given any such information. All participants then responded to our main repeated dependent measures: (1) actual guilt: “How guilty do you feel

about not going to the gym?" and (2) should guilt: "How guilty *should* you feel about not going to the gym" (1 = *not at all guilty*, 7 = *extremely guilty*).

Results

We conducted a linear mixed model with random by-participant intercepts using the lmerTest package (Kuznetsova, Brockhoff, & Christensen, 2014) in R statistical software (R Core Team, 2014), regressing guilt on guilt type (should vs. actual), overdetermined outcome condition, and their interaction[5], which yielded a significant interaction, $t(299) = 7.06$, $p < .001$ (see Figure 1). When the outcome was overdetermined (i.e., the gym was closed), people indicated that they should feel guiltier ($M = 3.54$, $SD = 1.85$) than they actually would feel ($M = 2.34$, $SD = 1.67$), $t(513.54) = 9.26$, $p < .001$, d = .68. When the outcome was not overdetermined (i.e., the gym was open), this effect was nonsignificantly reversed such that people thought they should feel *less* guilty ($M = 3.86$, $SD = 1.78$) than they actually would feel ($M = 4.22$, $SD = 1.76$), $t(513.54) = 1.56$, $p = .12$, d = .21. Decomposing the interaction differently, participants thought they *would* feel less guilty when they could point to their action being overdetermined than when they could not, $t(299) = –7.61$, $p < .001$, d = 1.09. This difference was significantly attenuated on the measure asking how they felt they *should* feel, but notably still significant, $t(299) = 2.36$, $p = .019$, d = .17. That is, participants acknowledged that the fact that their intended outcome (no exercise) was overdetermined *should not* have alleviated their guilt to the extent that it *did*.

[5] Note that mixed models often yield non-integer degrees of freedom.



Note. Error bars are 95% confidence intervals.

Figure 1. How guilty one should feel and how guilty one would feel, as a functions of overdetermined outcome condition

Discussion

In this study, we found that participants believe that the palliative effect of finding that the negative consequence of their intended actions was overdetermined is consolation overreach. That is, although people acknowledge they *would* feel better if they were to find that the negative outcome would have happened even without their weakness of will, they do not believe that they *should* feel better to the extent that they would (H3). Looked at differently, they appear to believe that, though they *should* feel somewhat better when their outcome was overdetermined, they were not justified in experiencing the degree of guilt reduction that they did.

Results from Studies 3 and 4 indicate that people both cognitive and motivated components play a role in evaluation of virtue. People believe that an overdetermined outcome—

that recycling was taken to the landfill (Study 3) or that the gym was closed (Study 4)—should rightfully alleviate guilt to some extent, as reflected in judgment of others (Study 3) and *ought* judgments of themselves (Study 4). Both studies suggest that people believe it is reasonable to feel less guilty for not acting virtuously (not recycling or not going to the gym) if their attempt to be virtuous (recycling or going to the gym) would have failed anyway. However, they believe their original intention—to not recycle or not go to the gym—should also be factored into their negative self-judgment. The results of Study 4 also indicate that people recognize that they underweight their sinful intentions when the outcome it leads to is overdetermined.

Studies 5A and 5B

The main goal of Studies 5A and 5B is to test H4: that people actively seek out evidence that an outcome is overdetermined when they feel guilt over a self-standard violation. In both studies, we manipulated whether participants imagined violating a self-control goal or not by assigning them to imagine choosing a cookie or a fruit salad. We expected that those who chose the cookie would be more likely than those who chose the fruit salad to incur an actual time cost to seek evidence of the non-consequential role played by their behavior (i.e., seeing if the cookie might have been stocked out). This would further support the motivated nature of the overdetermined outcome defense, as counterfactuals should be equally informative for both choices.

In Study 5A, we measured guilt and expected it to mediate this effect. In Study 5B, in addition to manipulating choice (violation vs. non-violation), we manipulated likelihood of the outcome's overdetermination: how likely the unchosen option was to be stocked out. As predicted by H5, we expected our effect to be robust to the probability of overdetermination, but explored linear and curvilinear effects of that probability on information seeking as well.

Method

Participants and Design

We opened Study 5A to 300 participants from Prolific, resulting in 301 participants (M_{age} = 301, 144 male, 154 female, 3 non-binary). This study had a two-condition (violation vs. non-violation choice) design. We opened Study 5B to 600 participants from Prolific, resulting in 599 participants (M_{age} = 41.60, 273 male, 315 female, 11 non-binary). This study had a 2 (choice: violation vs. non-violation) × continuous (probability of overdetermination: 10, 20, 30, 40, 50, 60, 70, 80, or 90 percent) design. Both studies were preregistered.

Procedure

In both studies, participants read the following, adapted from a scenario from Study 1: "Imagine that you have recently been trying to eat healthy. You are in line at a cafe, and want to order a snack. The cafe has a cookie and a fruit salad on the menu, and you are choosing between the two. You decide to go for the cookie (violation condition) / fruit salad (non-violation condition)." In Study 5A only, they indicated the extent to which they felt guilty about their decision (1 = *not at all*, 5 = *extremely*), which served as our mediator measure.

In Study 5A, participants then read, "The café often runs out of fruit salads (violation condition) / cookies (non-violation condition) and therefore you may not have been able to order one anyway." In Study 5B, participants instead read, "Based on their history, there is a **[1-9]0%** chance that the cafe ran out of fruit salads / cookies. Therefore, there is a **[1-9]0%** chance you would not have been able to order one anyway." Thus, chance of overdetermination was a random multiple of 10 between 10 and 90; as preregistered, we treated this manipulation as a continuous independent variable.

Participants in both studies then read, "You did not know whether the fruit salads / cookies were in-stock or out of stock when you chose to order the cookie / fruit salad. Do you want to find out now whether the cafe was stocked out of fruit salads / cookies? Note that you will have to wait 15 seconds to find out." Participants selected one of two options as our main information seeking measure: "Find out whether the fruit salads / cookies were stocked out (wait 15 seconds), or "Do not find out whether the fruit salads / cookies were stocked out."

Results

Main results for both studies. Binary logistic regressions indicated that participants in the violation condition in both studies were indeed more likely to check to see if the unordered item was stocked out than participants in the non-violation condition: Study 5A: 53.3% vs. 17.2%, B = 1.70, SE = .28, $p < .001$, OR = 5.49; Study 5B: 40.1% vs. 13.0%, B = 1.50, SE = .21, $p < .001$, OR = 4.49.

Study 5A: Mediation. In Study 5A, as expected, those who ordered the cookie felt guiltier ($M = 2.66$, $SD = 1.08$) than those who ordered the fruit ($M = 1.36$, $SD = .88$), $t(299) = 11.44$, $p < .001$. When added into the logistic regression predicting whether they would check if the unordered item was stocked out, the effect of guilt was significant, B = –.62, SE = .13, $p < .001$, and the effect of condition was reduced, B = 1.00, SE = .31, $p = .001$. A bootstrapping mediation analysis (5,000 sims) revealed that guilt significantly mediated the effect, 95% CI: [.09, .23].

Study 5B: Probability of Overdetermination. We ran a logistic regression regressing overdetermination information seeking behavior onto choice condition and linear probability of overdetermination. We found no interaction of the probability of overdetermination and choice condition, B = –.08, SE = .08, $p = .34$.

We next ran this analysis including the quadratic effect of probability of overdetermination. We found a significant interaction between choice condition and the quadratic effect of probability of overdetermination, B = 10.64, SE = 5.42, p = .049. The quadratic effect of probability on overdetermination information was nonsignificant in the non-violation condition, B = 5.91, SE = 4.59, p = .20, and marginal in the violation condition, B = –4.73, SE = 2.87, p = .099. Participants were most likely to check information when uncertainty was low (i.e., at low and high chance of overdetermination), and least likely to check when uncertainty was high (i.e., at moderate chance of overdetermination). Importantly, however, the main effect of violation condition was significant across each probability condition.

Discussion

In these studies, we found that people were more likely to indicate that they would check whether the outcome of their behavior was overdetermined if their behavior was a self-standard violation than if it was not (H4). This provides evidence of the motivated nature of this information seeking, as the counterfactual should rationally inform the amount of credit one is able to take for one's self-standard adherence just as much as it informs the amount of blame one is able to take for one's self-standard violations. Thus, a purely cognitive process would predict equal information seeking when one either violated or adhered to self-standards.

Further providing evidence of motivated information seeking, Study 5A showed that those who chose an unhealthy snack over a healthy one felt guiltier than those who chose a healthy snack over an unhealthy one, which led them to incur a material time cost to seek information about whether their decision was overdetermined—specifically, whether the healthy snack had been stocked out and that they therefore could not have ordered it anyway.

In Study 5B, we found this effect to be robust across different likelihoods of overdetermination. We also found suggestive evidence that people were least likely to check whether an outcome was overdetermined at the highest levels of uncertainty (when there was a moderate chance that their outcome was overdetermined). Although even at moderate levels of likelihood people sought this information more when they violated their personal standards than when they did not, this may reflect that some people take advantage of the uncertainty to preserve the possibility that the outcome was indeed overdetermined, rather than possibly finding out that it was not. Put differently, when the fruit was not likely to be stocked out, people may feel they have little to lose by checking—although finding out the fruit was stocked out would be a pleasant surprise, finding out that it was not stocked out would not lower their already low self-evaluation. On the other hand, when the fruit was likely to be stocked out, people may feel they have a lot to gain by checking—likely, they would confirm their hope that they could not have eaten the fruit anyway. However, when the fruit may or may not have been stocked out, people may already feel some form of appeasement of their guilt; some may have wanted to keep this hope alive and avoid finding out its falsehood. Future research is necessary to replicate this finding and further examine possible psychological mechanisms.

Study 6

In Study 6, we examined the overdetermined outcome defense in the context of prosocial giving. First, we sought to again test H1 by demonstrating that learning that a self-standard violation (i.e., choosing not to donate) was not necessary for the occurrence of the negative outcome (i.e., that the money would not have been donated anyway) would alleviate the actor's guilt. Furthermore, we examined the effect of people's experienced guilt on their overdetermined outcome information seeking behavior. Specifically, we tested whether people would actively

seek evidence that the consequence of their self-standard violation was overdetermined when they felt guilty, thus testing H4 in an incentive-compatible behavioral context.

Method

Participants and design

We preregistered this study[6] and opened it to 600 participants from Amazon's Mechanical Turk, resulting in 611 participants (M_{age} = 38.99, 327 male, 269 female, 3 non-binary, 12 did not respond). This study had a 2 (outcome: overdetermined vs. non-overdetermined control) × 2 (guilt: before vs. after outcome reveal) mixed-subjects design in which outcome condition was randomly assigned between subjects and guilt was measured within subject before and after the outcome (i.e., overdetermined or non-overdetermined) was revealed. Based on pilot data, we expected that recruiting 600 participants would result in at least 200 participants who chose to first keep the bonus (i.e., engage in a self-standard violation) and then also chose to view the outcome (i.e., whether the money would have been donated or not). Thus, we aimed to recruit at least 100 participants per outcome condition. In the end, 482 participants kept the bonus and 258 of those chose to view the outcome.

Procedure

Participants were told that they would be making an economic decision, and then given the option between two outcomes: receiving a $0.25 bonus payment or donating either $0 (with 50% probability) or $0.50 (with 50% probability) to the American Cancer Society. After making the decision, they indicated how guilty, ashamed, embarrassed, and proud they felt (the four

[6] Note one error in the pre-registration under #8 in which we reference a continuous measure that we had opted to remove from the study, and accurately removed from the dependent measure section, but neglected to remove from exploratory analyses.

major self-conscious emotions, Tangney, 2005), with a preregistered a priori decision to focus on guilt as our main dependent measure.

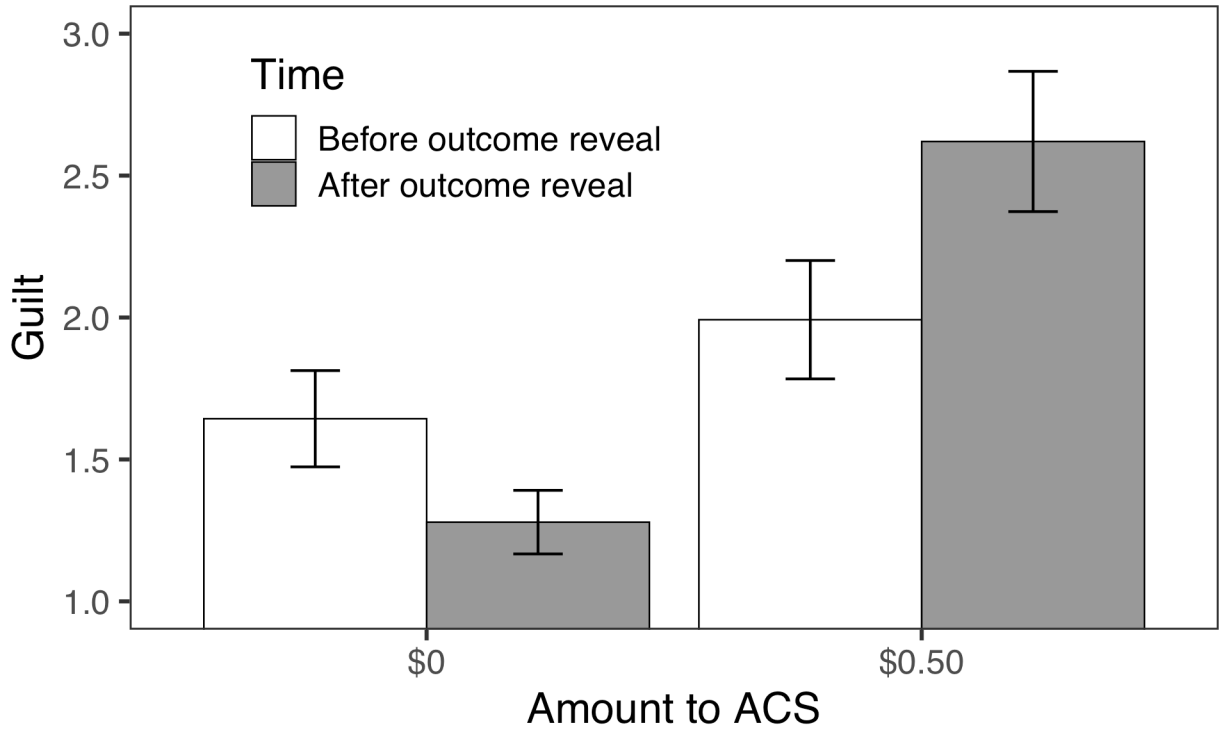
Those who chose to donate the money then completed the study. Those who chose to receive the bonus were directed to the rest of the study and constituted the sample of the study (n = 482). They were asked: “Would you like to find out whether the amount donated to the American Cancer Society would have been 0 or $.50? If you select "yes," you will be told whether the amount was $0 or $.50.” They indicated yes or no, which was our measure of overdetermined outcome seeking. That is, if participants were to find out that no money would have been donated, they could use this as a justification for their choice, even though they made the choice without knowing whether the amount would have been $0 or $0.50. Next, they indicated the extent to which the following reasons contributed to their decision to find out or not: “I didn’t care,” “I would feel better about my choice if no money would have been donated to ACS,” “I would feel worse about my choice if $.50 would have been donated to ACS,” and “I was curious” (1 = *not at all*, 5 = *extremely*).

Participants were then told that the amount to ACS would have been either $0 or $0.50 (randomly assigned). Finally, they indicated how much they were feeling the same self-conscious emotions that they did earlier. After the study was completed, bonuses were paid to those who chose the keep the bonus, and a $29.75 donation was made to ACS on behalf of the 119 participants who chose to donate.

Results

Among the 482 participants who chose the bonus, 53.5% (n = 258) chose to view whether any money would have been donated to ACS. First, we tested whether finding out that no donation would have been made decreased guilt for these participants. We employed a linear

mixed-model regressing guilt on amount that would have been donated to ACS ($0 vs. $.50) and time (before vs. after viewing outcome), where the amount to ACS was manipulated between subjects and guilt was measured over time within subject, including random intercepts for each participant. We found a significant effect of the interaction between amount to ACS and time on how guilty participants felt, $t(256) = 8.16$, $p < .001$ (see Figure 2). Participant's guilt was reduced when they found out that no donation would have been made ($M_{before} = 1.64$, $SD = .98$; $M_{after} = 1.28$, $SD = .65$), $t(269) = –4.24$, $p < .001$, d = .44. On the other hand, people felt guiltier than before when they found out that $0.50 would have been donated to ACS ($M_{before} = 1.99$, $SD = 1.21$; $M_{after} = 2.62$, $SD = 1.43$), $t(269) = 7.31$, $p < .001$, d = .47.



Note: Error bars are 95% confidence intervals.

Figure 2. Experienced guilt before and after finding out donation outcome as a function of donation outcome

Looked at differently, participants felt less guilty when they found out that no money would have been donated to ACS than if $0.50 would have been donated, $t(372.89) = 9.73$, $p < .001$, d = 1.21. Unexpectedly, participants in the overdetermined outcome (i.e., $0) condition felt less guilty than participants in the control (i.e., $0.50) condition before they found out the outcome, $t(372.89) = 2.53$, $p = .012$, d = .32, indicating the those in the overdetermined outcome group felt less guilty at baseline. However, the interaction indicates that the difference in guilt was significantly (i.e., one scale point) larger after viewing the outcome. Arguably, the fact that guilt differed by condition at baseline could be seen as creating a more conservative test of the hypothesis, as those in the overdetermined outcome condition already felt relatively lower guilt before viewing the outcome.

We next tested whether experienced guilt predicted whether participants viewed the outcome. We employed a binomial logistic regression in which we regressed choice to view the outcome onto measured guilt (before choice of whether to view whether the outcome was overdetermined). Results revealed that the guiltier participants felt about their choice the more likely they were to choose to view it, B = .45 SE = .10, $p < .001$, OR = 1.57. We then regressed guilt onto the various reasons they made their decisions (i.e., because it would make them feel better if money would not have been donated [“feel better”], because it would make them feel worse if money would have been donated [“feel worse”], because they were curious, and because they didn’t care). Guilt positively predicted the “feel better” item, $t(477) = 10.56$, $p < .001$, the “feel worse” item, $t(478) = 12.18$, $p < .001$, and curiosity, $t(479) = 5.41$, $p < .001$. Guilt also negatively predicted not caring, $t(479) = –3.99$, $p < .001$.

We then ran a logistic regression in which we regressed whether participants chose to view or not view the donation outcome onto guilt along with the “feel better,” “feel worse,” and

"I didn't care" variables. As preregistered, we did not include the curiosity measure in the regression because curiosity may not explain unique variance from the "feel better" variable; that is, people may be curious because they would feel better if they found out their outcome was overdetermined. We found that the "feel better" item was a significant predictor, B = .39, SE = .10, $p < .001$, and guilt was no longer a predictor, B = .11, SE = .13, $p = .39$. The "feel worse" item also did not predict information seeking, B =.05, SE = .11, $p = .65$. The "didn't care" item negatively predicted information seeking, B = –.78, SE = .09, $p < .001$. Furthermore, a bootstrapped mediation analysis (5,000 sims) found that the effect of guilt on the desire to view to the donation outcome was mediated by the "feel better" item, 95% CI: [–.04, –.01] (controlling for both the "feel worse" item and "didn't care" items).

Discussion

In a behavioral setting with real stakes, we found that participants felt less guilty when they learned that the charity they refused to donate to would not have benefitted even if they had originally decided to donate. That is, among those who chose to keep a bonus over donating the money, those who learned that a donation would not have been made anyway experienced diminished guilt, whereas those who learned that a donation would have been made experienced increased guilt.

As before, finding out that the outcome was overdetermined reduced guilt (even though they knew they acted without such foresight), providing further evidence for H1. Notably, in this study, the magnitude of increased guilt in response to the overdetermined outcome was larger than the magnitude of diminished guilt in the non-overdetermined outcome; this stands in contrast to Study 2, in which negative self-judgment was alleviated by an overdetermined outcome to a greater extent than it was enhanced in response to the non-overdetermined

outcome. However, post-hoc analyses revealed that people predicted the opposite: people agreed more in general with the "feel better" item (i.e., "I would feel better about my choice if no money would have been donated to ACS"; $M = 2.30$, $SD = 1.41$) than the "feel worse" item (i.e., "I would feel worse about my choice if $.50 would have been donated to ACS"; $M = 2.08$, $SD = 1.26$), $t(946.22) = 2.61$, $p = .009$. This indicates that, when choosing whether to view the outcome or not, participants underestimated the enhanced guilt they would have felt if they were to find out the donation would have occurred. Post-hoc analyses further provided evidence that the desire to avoid feeling worse did not play a role in information avoidance.[7]

Importantly, we found that those who felt guiltier about their self-standard violation were more likely to check whether the outcome of their action was negative or not. Specifically, the guiltier that people felt about keeping money for themselves rather than donating it, the more likely they were to check whether the money would have been donated or not. Mediation analysis suggests that guilt drove them to seek out information that could make them feel better; this drive dominated any potential fear that the information might confirm the harmfulness of their behavior (i.e., that the amount donated would have been higher than the amount they received).

General Discussion

Across seven studies we document a novel phenomenon we call the overdetermined outcome defense. People feel less guilty when their intention to violate a self-standard was not necessary for the occurrence of the negative consequences (Studies 1, 2, 4, and 6). Although

[7] The "feel worse" and "feel better" items were positively correlated, r = .53. The "feel worse" item (i.e., "I would feel worse about my choice if $.50 would have been donated to ACS") *positively* (rather than negatively) predicted information seeking when predicting information seeking on its own, B = .45, SE = .08, $p < .001$; when included with the "feel better" item, the "feel better" item significantly predicted information seeking, B = .50, SE = .09, $p < .001$, whereas the "feel worse" item was reduced, B = .18, SE = .09, $p = .062$.

their intended self-standard violation was sufficient for the occurrence of the negative consequences, they felt less guilty when their intentional acts did not matter in the end. This effect is at least partially motivated (Studies 3 and 4). Unlike other justification techniques, people seem aware that their reliance on this tactic is illegitimate (Study 4). They acknowledge that the guilt they feel over a negative overdetermined outcome reflects their intentionality less than it should. The guilt reduction that they experience when they can invoke the overdetermined outcome defense drives them to seek information in search of such overdetermined outcomes (Studies 5A, 5B, 6). Furthermore, they seek such information regardless of the likelihood that the outcome is over determined (Study 5B).

Theoretical Contributions

Previous research on how people justify engaging in self-standard violations or engaging in *want* over *should* behaviors in the moral or self-control domains, has pointed to the psychological interpretations people impose on their behavior that serve to minimize the threat to their moral or virtuous self-concept (Huberts et al., 2014; Shalvi et al., 2015). These interpretations include decreasing the perceived harmfulness of one's initial intentions or bolstering one's own perceived virtuousness such that one is licensed to engage in such a transgression without harming one's self-view. In contrast, we find that people can use, and indeed actively search for, an element of the decision that is exogenous to both their intentions and to themselves—the chance outcomes of their actions—to justify their behavior. This complements prior research showing that people can use risk as a justification for self-standard violations, such as not donating (e.g., people will not donate if there is a chance their donation outcome will not be realized; Exley, 2015). In this prior research, the fact that an outcome might be overdetermined justifies people's decisions to violate their self-standards. In contrast, rather

than using an overdetermined circumstance to justify their decisions *a priori*, we demonstrate that people actively seek out these circumstances *post hoc* to retroactively justify their decision.

Relatedly, our research contributes to the literature on outcome bias (Baron & Hershey, 1988; Gino, Shu, & Bazerman, 2010). Instead of considering how observers judge decisions differently depending on the outcomes they produced, we examine how actors judge decisions differently depending on whether the outcome would have occurred independently of one's decision. Our research highlights that individuals will engage in post-hoc justification processes even when they recognize the extent to which they have violated their self-standards. Furthermore, we find that people are willing to spend time to ascertain the presence of overdetermined outcomes, even though they are aware such information seeking is at only somewhat legitimate.

One important point to note in comparing the previous work on outcome bias with the present work on overdetermined outcomes is the way in which these two situations are treated differently in the legal domain. Whereas outcome bias, at least in certain circumstances, is enshrined in law (e.g., attempted murder necessitates a lesser sentence than successful murder), overdetermined outcomes provide a much more complicated issue for legal determinations of responsibility. A case where a group of people commits a crime together does not diminish the culpability of each individual perpetrator (Zimmerman, 1985); however, most current legal frameworks still struggle with cases of overdetermination (Lagnado, Gerstenberg, & Zultan, 2013). While the present paper cannot provide guidance on how to normatively assign blame in overdetermined situations, it does provide descriptive evidence that people use these opportunities to absolve themselves of guilt. Future research should examine how these justification strategies may impact attributions of blame in the legal setting.

This type of behavior is also related to work on quasi-magical thinking (Shafir & Tversky, 1992) and acquiescence (Risen, 2015), which involve acting on superstitious or false beliefs that one rationally knows should not be true. However, existing work in this domain has primarily examined contexts in which these beliefs are not only unhelpful to the decision maker but are often actively costly. For example, people incorporate their intuitions about probability into decisions in lotteries and games of chance (even though they know these intuitions are wrong), which leads them to lose money compared to making the more "rational" choice (Walco & Risen, 2017). In contrast, the present work presents an instance in which these "knowingly unreasonable" decisions are used to one's advantage. As a result, this represents some of the first evidence of people strategically using what they admit is irrelevant information to enhance their own self-image.

Furthermore, we find that the use of overdetermined outcomes as a defense strategy has both cognitive and motivated components. Although overdetermined outcomes lead to less negative judgments of others' sinful decisions (a cognitive process), they do so more for one's own decisions (a motivated process). Additionally, people believe that overdetermined outcomes should alleviate guilt to an extent (a cognitive process)—just not to the extent that they actually alleviate guilt (a motivated process).

The present findings speak to the flexibility of people's use of moral frameworks. Research indicates that people view non-consequentialist reasoning, such as deontological or person-centered judgments, to be more rooted in morality than consequentialist reasoning, and subsequently more relevant in judgment of moral character (Kreps & Monin, 2014; Uhlmann, Pizarro, & Diermeier, 2015). This suggests that people tend to believe that the determinants of moral character and culpability depend largely on a person's core values and intentions (Bertram

F Malle & Knobe, 1997; Bertram F Malle, Moses, & Baldwin, 2001) which leads them to denigrate those who ascribe to a consequentialist moral framework (Everett, Faber, Savulescu, & Crockett, 2018). Our research suggests an exception to this pattern: People will overweigh consequences when those consequences happen to exonerate themselves from blame. As such, people seem to engage in motivated shifting of their use of frameworks depending on whether it serves their self-judgment—that virtuousness is defined by how pivotal their action was in producing the outcome, rather than by the intention that lies behind their action. Perhaps it is because this shift in thinking is too large to be wholly self-deceptive that people are aware of this biased behavior.

Limitations and Future Directions

In this research we treat self-standard violations as similar in the moral and self-control domains. These two domains indeed share many similarities: *should* choices require high self-control (Baumeister & Exline, 1999; Gino, Schweitzer, Mead, & Ariely, 2011; Hofmann, Baumeister, Förster, & Vohs, 2012), and *want* choices cause guilty self-reproach (Baumeister et al., 1994; Dahl et al., 2003) in both domains. Also in both domains, guilt leads to licensing and justification strategies that protect one's self-view (Cascio & Plant, 2014; Effron, Monin, & Miller, 2013; Khan & Dhar, 2006, 2007; May & Irmak, 2014; Monin & Miller, 2001; Shalvi et al., 2015). In our research, we find that intentions are similarly disregarded in moral (Studies 1, 3 and 6) and self-control (Studies 1, 2, 4, 5A and 5B) domains when overdetermined outcomes negate the effects of those intentions. However, future research might examine whether the overdetermined outcome defense is considered more "reasonable" in the self-control domain than in the moral domain, as outcomes (rather than intentions) may matter more in those domains

(Berman & Small, 2018); these outcomes may then be sought more frequently, and may alleviate guilt to an even higher extent, than similar outcomes in the moral domain.

Future research may also investigate how self-judgments differ after finding out whether one's outcome was overdetermined or not among those who actively seek overdetermined information and for those who avoid it. Our studies showed that, on average, finding out one's negative outcome was overdetermined alleviated negative self-judgment for both those who would have sought it or not (Studies 1-4), and additionally that many do seek overdetermination information (Studies 5-6). However, those who actively avoided it may do so because they anticipate different affective reactions to that information. Although Study 6 did not find that anticipated guilt for non-overdetermined outcomes predicted information avoidance, it could be that those who avoid overdetermination information on average do not feel appeased by learning that their outcomes were overdetermined.

Future research might also examine the extent to which overdetermined outcomes can facilitate or impede perceived goal progress. If someone skips the gym, but then finds out the gym happened to be closed, she might be more likely to try to make up for the lost day later because she can externalize the blame for the initial decision not to go, maintaining a sense of herself as a gym-goer. Alternatively, she may be less likely to go at a later time, since the overdetermined outcome alleviates the guilt that could otherwise have been the impetus to make up for the lost day.

Finally, future work should examine the extent to which these findings are culturally bounded. In particular, the causes and consequences of feeling guilty differ by culture (Wong & Tsai, 2007), which may impact the overdetermined outcome defense. For example, previous work has found that in collectivist cultures such as Japan, negative self-evaluations are seen as

more meaningful and important to one's sense of self than in more individualistic such as the United States (Kitayama, Markus, Matsumoto, & Norasakkunkit, 1997). As a result, people in collectivist cultures may feel less of a need to assuage their guilt, and thus may be less likely to seek out overdetermined outcomes.

Concluding Remarks

We identify and document the overdetermined outcome defense. Our research shows that people feel less bad about themselves following an intentional self-standard violation when their action was not necessary for the occurrence of the outcome: “it wouldn’t have mattered anyway.” It also documents people’s motivation to seek out and incorporate information about “what would have been” to reduce their guilt for their sinful actions. As such, this research contributes to work on justification strategies, outcome bias, and quasi-magical thinking.

References

Agrawal, N., & Maheswaran, D. (2005). Motivated reasoning in outcome-bias effects. *Journal of Consumer Research*, *31*, 798–805.

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*, 364–374.

Baron, J., & Hershey, J. C. (1988). Outcome bias in decision evaluation. *Journal of Personality and Social Psychology*, *54*, 569–579.

Baumeister, R. F., & Exline, J. J. (1999). Virtue, personality, and social relations: Self-control as the moral muscle. *Journal of Personality*, *67*, 1165–1194.

Baumeister, R. F., & Heatherton, T. F. (1996). Self-regulation failure: An overview. *Psychological Inquiry*, *7*, 1–15.

Baumeister, R. F., Stillwell, A. M., & Heatherton, T. F. (1994). Guilt: An interpersonal approach. *Psychological Bulletin*, *115*, 243–267.

Berman, J. Z., & Small, D. A. (2012). Self-interest without selfishness: The hedonic benefit of imposed self-interest. *Psychological Science*, *23*, 1193–1199.

Berman, J. Z., & Small, D. A. (2018). Discipline and desire: On the relative importance of willpower and purity in signaling virtue. *Journal of Experimental Social Psychology*, *76*, 220–230.

Bitterly, T. B., Mislavsky, R., Dai, H., & Milkman, K. L. (2015). Want/should conflict: A synthesis of past research. In W. Hofmann & L. Nordgren (Eds.), *The Psychology of Desire* (pp. 244–264). New York, NY: The Guilford Press.

Cascio, J., & Plant, E. A. (2014). Prospective moral licensing: Does anticipating doing good later allow you to be bad now? *Journal of Experimental Social Psychology*, *56*, 1–8.

Dahl, D. W., Honea, H., & Manchanda, R. V. (2003). The nature of self-reported guilt in consumption contexts. *Marketing Letters*, *14*, 159–171.

DeMarree, K. G., Wheeler, S. C., Briñol, P., & Petty, R. E. (2014). Wanting other attitudes: Actual–desired attitude discrepancies predict feelings of ambivalence and ambivalence consequences. *Journal of Experimental Social Psychology*, *53*, 5–18.

Effron, D. A., Miller, D. T., & Monin, B. (2012). Inventing racist roads not taken: The licensing effect of immoral counterfactual behaviors. *Journal of Personality and Social Psychology*, *103*, 916–932.

Effron, D. A., Monin, B., & Miller, D. T. (2013). The unhealthy road not taken: Licensing indulgence by exaggerating counterfactual sins. *Journal of Experimental Social Psychology*, *49*, 573–578.

Everett, J. A. C., Faber, N. S., Savulescu, J., & Crockett, M. J. (2018). The costs of being consequentialist: Social inference from instrumental harm and impartial beneficence. *Journal of Experimental Social Psychology*, *79*, 200–216.

Exley, C. L. (2015). Excusing selfishness in charitable giving: The role of risk. *The Review of Economic Studies*, *83*, 587–628.

Gino, F., Schweitzer, M. E., Mead, N. L., & Ariely, D. (2011). Unable to resist temptation: How self-control depletion promotes unethical behavior. *Organizational Behavior and Human Decision Processes*, *115*, 191–203.

Gino, F., Shu, L. L., & Bazerman, M. H. (2010). Nameless + harmless = blameless: When seemingly irrelevant factors influence judgment of (un)ethical behavior. *Organizational Behavior and Human Decision Processes*, *111*, 93–101.

Golman, R., Hagmann, D., & Loewenstein, G. (2017). Information Avoidance. *Journal of Economic Literature*, *55*, 96–135.

Hofmann, W., Baumeister, R. F., Förster, G., & Vohs, K. D. (2012). Everyday temptations: An experience sampling study of desire, conflict, and self-control. *Journal of Personality and Social Psychology*, *102*, 1318–1335.

Huberts, J. C. D. W., Evers, C., & Ridder, D. T. D. de. (2014). "Because I am worth it": A theoretical framework and empirical review of a justification-based account of self-regulation failure. *Personality and Social Psychology Review*, *18*, 119–138.

Johnson, E. J., Häubl, G., & Keinan, A. (2007). Aspects of Endowment: A Query Theory of Value Construction. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*, 461–474.

Jonas, E., Schulz-Hardt, S., Frey, D., & Thelen, N. (2001). Confirmation bias in sequential information search after preliminary decisions: An expansion of dissonance theoretical research on selective exposure to information. *Journal of Personality and Social Psychology*, *80*, 557–571.

Kahneman, D., & Miller, D. T. (1986). Norm theory: Comparing reality to its alternatives. *Psychological Review*, *93*, 136–153.

Kelley, H. H. (1973). The processes of causal attribution. *American Psychologist*, *28*, 107–128.

Khan, U., & Dhar, R. (2006). Licensing effect in consumer choice. *Journal of Marketing Research*, *43*, 259–266.

Khan, U., & Dhar, R. (2007). Where there is a way, is there a will? The effect of future choices on self-control. *Journal of Experimental Psychology: General*, *136*, 277–288.

Khan, U., Dhar, R., & Wertenbroch, K. (2005). *A behavioral decision theoretic perspective on hedonic and utilitarian choice* (S. Ratneshwar & D. G. Mick, Eds.). In (pp. 144-165.).

Kitayama, S., Markus, H. R., Matsumoto, H., & Norasakkunkit, V. (1997). Individual and collective processes in the construction of the self: Self-enhancement in the United States and self-criticism in Japan. *Journal of Personality and Social Psychology*, *72*, 1245–1267.

Kivetz, R., & Zheng, Y. (2006). Determinants of justification and self-control. *Journal of Experimental Psychology: General*, *135*, 572–587.

Klein, N., & Epley, N. (2016). Maybe holier, but definitely less evil, than you: Bounded self-righteousness in social judgment. *Journal of Personality and Social Psychology*, *110*, 660–674.

Kreps, T. A., & Monin, B. (2014). Core values versus common sense: Consequentialist views appear less rooted in morality. *Personality and Social Psychology Bulletin*, *40*, 1529–1542.

Kruglanski, A. W. (1980). Lay epistemo-logic—process and contents: Another look at attribution theory. *Psychological Review*, *87*, 70–87.

Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin*, *108*, 480–498.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2014). lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, *82*, 1–26.

Lagnado, D. A., Gerstenberg, T., & Zultan, R. (2013). Causal Responsibility and Counterfactuals. *Cognitive Science*, *37*, 1036–1073.

Lin, S. C., & Miller, D. T. (2021). A dynamic perspective on moral choice: Revisiting moral hypocrisy. *Organizational Behavior and Human Decision Processes*, *164*, 203–217.

Lin, S. C., Zlatev, J. J., & Miller, D. T. (2017). Moral traps: When self-serving attributions backfire in prosocial behavior. *Journal of Experimental Social Psychology*, *70*, 198–203.

Liu, P. J., & Lin, S. C. (2018). Projecting lower competence to maintain moral warmth in the avoidance of prosocial requests. *Journal of Consumer Psychology*, *28*, 23–39.

Malle, Bertram F, & Knobe, J. (1997). The folk concept of intentionality. *Journal of Experimental Social Psychology*, *33*, 101–121.

Malle, Bertram F, Moses, L. J., & Baldwin, D. A. (2001). *Intentions and Intentionality*. MIT Press.

May, F., & Irmak, C. (2014). Licensing indulgence in the present by distorting memories of past behavior. *Journal of Consumer Research*, *41*, 624–641.

McFarland, C., & Ross, M. (1982). Impact of causal attributions on affective reactions to success and failure. *Journal of Personality and Social Psychology*, *43*, 937–946.

Meyer, J. P. (1980). Causal attribution for success and failure: A multivariate investigation of dimensionality, formation, and consequences. *Journal of Personality and Social Psychology*, *38*, 704–718.

Milkman, K. L., Rogers, T., & Bazerman, M. H. (2008). Harnessing our inner angels and demons: What we have learned about want/should conflicts and how that knowledge can help us reduce short-sighted decision making. *Perspectives on Psychological Science*, *3*, 324–338.

Miller, D. T., Turnbull, W., & McFarland, C. (1990). Counterfactual thinking and social perception: Thinking about what might have been. *Advances in Experimental Social Psychology*, *23*, 305–331.

Monin, B., & Miller, D. T. (2001). Moral credentials and the expression of prejudice. *Journal of Personality and Social Psychology*, *81*, 33–43.

Norton, M. I., Vandello, J. A., & Darley, J. M. (2004). Casuistry and social category bias. *Journal of Personality and Social Psychology*, *87*, 817–831.

Okada, E. M. (2005). Justification effects on consumer choice of hedonic and utilitarian goods. *Journal of Marketing Research*, *42*, 43–53.

Pearl, J. (2009). *Causality* (2nd ed.). Cambridge University Press.

Pronin, E., Gilovich, T., & Ross, L. (2004). Objectivity in the eye of the beholder: Divergent perceptions of bias in self versus others. *Psychological Review*, *111*, 781–799.

Pronin, E., Lin, D. Y., & Ross, L. (2002). The bias blind spot: Perceptions of bias in self versus others. *Personality and Social Psychology Bulletin*, *28*, 369–381.

Reczek, R. W., Irwin, J. R., Zane, D. M., & Ehrich, K. R. (2017). That's not how I remember it: Willfully ignorant memory for ethical product attribute information. *Journal of Consumer Research*, *45*, 185–207.

Risen, J. L. (2015). Believing what we do not believe: Acquiescence to superstitious beliefs and other powerful intuitions. *Psychological Review*, *123*, 182–207.

Roese, N. J. (1997). Counterfactual thinking. *Psychological Bulletin*, *121*, 133–148.

Schulz-Hardt, S., Frey, D., Lüthgens, C., & Moscovici, S. (2000). Biased information search in group decision making. *Journal of Personality and Social Psychology*, *78*, 655–669.

Scopelliti, I., Morewedge, C. K., McCormick, E., Min, H. L., Lebrecht, S., & Kassam, K. S. (2015). Bias blind spot: Structure, measurement, and consequences. *Management Science*, *61*, 2468–2486.

Shafir, E., & Tversky, A. (1992). Thinking through uncertainty: Nonconsequential reasoning and choice. *Cognitive Psychology*, *24*, 449–474.

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

Shaver, K. G. (1970). Defensive attribution: Effects of severity and relevance on the responsibility assigned for an accident. *Journal of Personality and Social Psychology*, *14*, 101–113.

Tangney, J. P., Stuewig, J., & Mashek, D. J. (2007). Moral emotions and moral behavior. *Annual Review of Psychology*, *58*, 345–372.

Team, R. C. (2014). *R: A Language and Environment for Statistical Computing*. Retrieved from http://www.R-project.org/

Uhlmann, E. L., & Cohen, G. L. (2005). Constructed criteria: Redefining merit to justify discrimination. *Psychological Science*.

Uhlmann, E. L., Pizarro, D. A., & Diermeier, D. (2015). A person-centered approach to moral judgment. *Perspectives on Psychological Science*, *10*, 72–81.

Walco, D. K., & Risen, J. L. (2017). The empirical case for acquiescing to intuition. *Psychological Science*, *28*, 1807–1820.

Weiner, B. (1985). An attributional theory of achievement motivation and emotion. *Psychological Review*, *92*, 548–573.

Wertenbroch, K. (1998). Consumption self-control by rationing purchase quantities of virtue and vice. *Marketing Science*, *17*, 317–337.

Wiltermuth, S. S. (2011). Cheating more when the spoils are split. *Organizational Behavior and Human Decision Processes*, *115*, 157–168.

Wong, Y., & Tsai, J. (2007). Cultural models of shame and guilt. In J. L. Tracy, R. W. Robins, & J. P. Tangney (Eds.), *The self-conscious emotions: Theory and Research* (pp. 209–223). Guilford Press.

Xu, J., & Schwarz, N. (2009). Do we really need a reason to indulge? *Journal of Marketing Research*, *46*, 25–36.

Zimmerman. (1985). Sharing responsibility. *American Philosophical Quarterly*, *22*, 115–122.

Zlatev, J. J., Kupor, D. M., Laurin, K., & Miller, D. T. (2020). Being "good" or "good enough": Prosocial risk and the structure of moral self-regard. *Journal of Personality and Social Psychology*, *118*, 242–253.

Zlatev, J. J., & Miller, D. T. (2016). Selfishly benevolent or benevolently selfish: When self-interest undermines versus promotes prosocial behavior. *Organizational Behavior and Human Decision Processes*, *137*, 112–122.

Zuckerman, M. (1979). Attribution of success and failure revisited, or: The motivational bias is alive and well in attribution theory. *Journal of Personality*, *47*, 245–287.