ANOTHER WRONGDOER’S PUNISHMENT CLEANSES THE SELF: EVIDENCE FOR A MORAL CLEANSING FUNCTION OF PUNISHING MORAL TRANSGRESSORS

BY

© 2013

ZACHARY KLEIN ROTHSCILD

Submitted to the graduate degree program in Psychology and the Graduate Faculty of the University of Kansas in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

_______________________________________
Chairperson Mark Landau, Ph.D.

_______________________________________
Christian Crandall, Ph.D.

_______________________________________
Nyla Branscombe, Ph.D.

_______________________________________
Ludwin Molina, Ph.D.

_______________________________________
Robert Antonio, Ph.D.

Date Defended: May 31, 2013
The Dissertation Committee for ZACHARY KLEIN ROTHSCCHILD certifies that this is the approved version of the following dissertation:

ANOTHER WRONGDOER’S PUNISHMENT CLEANSES THE SELF: EVIDENCE FOR A MORAL CLEANSING FUNCTION OF PUNISHING MORAL TRANSGRESSORS

________________________________

Chairperson Mark Landau, Ph.D.

Date approved: June 6, 2013
ABSTRACT

Separate lines of research show that individuals: a) understand immorality metaphorically as physical contamination; b) project undesirable self-attributes onto others; and c) view punishment as eliminating a transgressor’s immorality. Integrating these findings, we hypothesized that individuals project guilt over their own immoral actions – experienced as physical contamination – onto another transgressor whose punishment restores their own perceived moral and physical purity. Study 1 showed that priming participants’ own immoral actions decreased their felt physical cleanliness, but not if they were presented with a punished transgressor. Study 2 showed that participants primed with their own immoral actions viewed another transgressor as physically dirtier as a result of their own increased feelings of guilt. Additionally, the subsequent punishment of the contaminated transgressor restored participants’ perceived personal morality and eliminated their desire to engage in moral restoration behavior. These studies are the first to show that another’s punishment can serve to “cleanse” the self.
Acknowledgements

I would like to thank Mark Landau, not only for his invaluable advice on this projection, but for his continuous support and guidance throughout my graduate education. I have benefited from him as a teacher, researcher, colleague and friend. All of my academic endeavors will, from this point forward, bear the mark of Mark. I would also like to thank Nyla Branscombe, Ludwin Molina, Chris Crandall, Monica Biernat, Glen Adams and Omri Gillath. Each of these exceptional people has helped to shape my perspective as a psychologist. My tribute to them will be to continue to pass on their wisdom to future generations of students. Finally, I would like to thank my fellow graduate students, in particular Daniel Sullivan and Lucas Keefer, without whom I would surely still be wandering the halls of Fraser, lost and confused.
Table of Contents

Title Page ........................................................................................................................................i
Acceptance Page ............................................................................................................................ii
Abstract ...........................................................................................................................................iii
Acknowledgements ........................................................................................................................iv
Table of Contents .............................................................................................................................v
List of Tables & Figures ......................................................................................................................vi
General Introduction .......................................................................................................................1
Study 1 ...............................................................................................................................................11
Method .............................................................................................................................................12
Results .............................................................................................................................................14
Discussion .......................................................................................................................................15
Supplemental Study ..........................................................................................................................16
Study 2 ...............................................................................................................................................17
Method .............................................................................................................................................19
Results .............................................................................................................................................22
Discussion .......................................................................................................................................26
General Discussion ...........................................................................................................................27
References .........................................................................................................................................40
Footnotes ...........................................................................................................................................45
Tables ...............................................................................................................................................46
Figures ...............................................................................................................................................48
Appendices .......................................................................................................................................52
List of Tables & Figures

Table 1: Ratings of target as “evil” and “dirty” as a function of target type, Supplemental Study........................................................................................................................................46

Table 2: Reported interest in joining blood drive as a function of personal immorality salience and target type, Study 2..................................................................................................................47

Figure 1: Ratings of personal physical cleanliness as a function of personal immorality salience and target type, Study 1.......................................................................................................................................48

Figure 2: Indirect effect of personal immorality salience on perceived physical dirtiness of target transgressor through feelings of personal guilt, Study 2..........................................................................................49

Figure 3: Ratings of personal morality relative to other students as a function of personal immorality salience and target type, Study 2........................................................................................................50

Figure 4: Odds ratios of reported interest in joining a blood drive as a function of personal immorality salience and target type, Study 2........................................................................................................51
Another Wrongdoer’s Punishment Cleanses the Self: Evidence for a Moral Cleansing Function of Punishing Moral Transgressors

Members of many pre-modern cultures regularly participated in “rituals of purification” in which a target individual or non-human animal was punished to restore the community’s moral purity (Victor, 2003). For instance, ancient Israelites placed their sins onto a goat that was exiled to expunge their guilt (Allport, 1954/1979), while the Greeks stoned criminals burdened with the transgressions of others in order to remove the community’s immorality (Frazer, 1922/2002). Although these purification practices differ in superficial respects, they are all based on three interrelated beliefs: immorality exists as a tangible contaminant; it can be transferred from the self to others; and it can be eliminated by punishing those “contaminated” others (Douglas, 1995).

Although these beliefs may sound archaic to modern ears, separate lines of research show that they continue to shape people’s thinking about morality: people understand immorality metaphorically as a tangible contaminant (Rozin, Lowery, & Ebert, 1994); they project undesired aspects of the self onto others (Schimel, Greenberg, & Martens, 2003); and they view punishment as a means of expiating the moral transgressions of the punished (Bastian, Jetten & Fasoli, 2011). Integrating and extending these findings, we hypothesized that individuals continue to believe that the punishment of other transgressors can serve to atone for their own immoral actions.

The current studies are the first to assess this hypothesis experimentally. Study 1 tests whether a salient personal immoral action leads individuals to feel physically dirtier unless they are exposed to a punished transgressor. Study 2 tests whether individuals’ project felt contamination elicited by their own salient immoral behavior onto another transgressor, and
whether the subsequent punishment of that transgressor restores their own positive moral identity. Study 2 also tests whether exposure to a punished transgressor eliminates individuals’ desire to engage in compensatory moral behavior in response to a salient personal transgression.

To frame these studies, below we provide more detail on the aforementioned lines of prior research; then we discuss how the current studies integrate and extend previous findings while providing evidence of a novel moral cleansing phenomenon.

**Overview of Previous Research**

**Moral and Physical Purity**

According to conceptual metaphor theory (CMT; Lakoff & Johnson, 1980), people use metaphor to understand abstract concepts in terms of dissimilar concepts which are more concrete and easier to understand. A conceptual metaphor creates a mapping between corresponding elements of the two concepts, which allows people to use knowledge of the concrete concept as a framework for interpreting and evaluating analogous elements of the abstract concept. This perspective suggests that common expressions like “dirty deed” and “clean conscience” are more than ornamental figures of speech; instead, they reflect an underlying conceptual metaphor that people use to understand *morality* in terms of *physical cleanliness* (Lakoff & Johnson, 1999).

Empirical evidence for this claim comes from recent studies showing that manipulating people’s experience of physical cleanliness affects their moral judgments in a metaphor-consistent manner, even though they are unaware of the influence of their physical surroundings. In one such series of studies (Schnall, Haidt, Clore, & Jordan, 2008), participants exposed to foul odor (Study 1) and participants who were seated in an unclean physical environment (Study 2) judged a variety of moral dilemmas (e.g., keeping money inside a found wallet, putting false
information on a resume’) to be more immoral. In a related study examining cleansing acts, participants who washed their hands rated these moral dilemmas as less wrong than those not given the opportunity to physically cleanse themselves (Schnall, Benton & Harvey, 2008).

Focusing on self-perceptions, Zhong, Strejcek, and Sivanathan (2010) showed that increasing participants’ perceptions of their own physical cleanliness led to increased ratings of their personal moral character. These studies indicate that the abstract thought of morality is grounded in and influenced by people’s concrete experience of physical cleanliness.

Research has also shown that perceptions of morality have reciprocal effects on perceptions of physical cleanliness. Studies show that moral violations can elicit a disgust response, commonly evoked by unclean physical environments and associated with the avoidance and expulsion of physical contaminants (Chapman, Kim, Susskind & Anderson, 2009; Haidt, Koller, & Dias, 1993; Moll et al., 2005; Rozin et al., 1994). Related findings show that people avoid contact with objects that have been previously owned by, or been in contact with, immoral people (Rozin, Markwith, & McCauley, 1994). Specifically, Rozen and colleagues found that participants were less willing to use an object (e.g., a sweater, a bed, an automobile) if it was previously owned by a murderer. The authors explain this effect as an attempt to avoid “moral taint.” In line with CMT, this finding illustrates how people use their knowledge about the spread of physical contaminants to judge the impact of actions on their own moral identity.

While the aforementioned research focuses on the avoidance of external sources of contamination, scholars have suggested that individuals can be “stained” by their own moral transgressions (e.g., Ricoeur, 1967). Recent empirical studies support the contention that individuals can feel physically soiled by their own immoral actions (Lee & Schwarz, 2010; Zhong & Liljenquist, 2006). Zhong and Liljenquist showed that participants reminded of their
own unethical (vs. ethical) behavior showed a greater desire to engage in physical cleansing behavior. Specifically, they found that participants whose own immorality was made salient were more likely to choose an antiseptic hand wipe over a pencil when offered both products as a gift for participating in the study. In another study, Zhong and Liljenquist manipulated whether or not participants physically cleansed themselves with an antiseptic wipe after being reminded of a personal moral transgression. They found that participants who contemplated their own immoral (vs. moral) actions reported increased feelings of guilt and an increased desire to engage in moral restoration behavior (e.g., volunteering to help another student in need), but both of these effects were eliminated if participants physically cleansed themselves.

These findings illustrate an embodied cognitive process in which one’s own salient wrongdoings elicit feelings of guilt which are understood as both a moral and physical contamination of the self. As such, the act of cleansing the self of physical contamination removes feelings of guilt and circumvents the desire to restore one’s perceived moral identity through other means. Of course, insofar as the relationship between morality and physical cleanliness is bidirectional (Lee & Schwarz, 2012), we would also expect motivated efforts to expel or deny one’s immorality to both physically and morally “cleanse” the self. In the next section we review research on defensive projection, a motivated mechanism associated with the expulsion and denial of undesirable aspects of the self.

**Projection**

Defensive projection is commonly defined as “the process of perceiving one’s own undesirable qualities in others as a way of protecting one’s self-image” (Govorun, Fuegen, & Payne, 2006, p. 781). Sigmund Freud (1915/1957) first described projection as a mechanism in which the ego "expels whatever within itself becomes a cause of displeasure" (p.136). Refining
her father’s theorizing, Anna Freud (1936) proposed what is now considered the classic account of projection. According to her formulation, when people are confronted with their faults, they are motivated to deny them by seeing those faults in others. In this account, projection is an unconscious defensive mechanism that serves to alleviate an individual’s anxiety by facilitating the denial of his or her own unacceptable thoughts, desires, and impulses, ultimately maintaining a positive self-image (see also Jung’s analysis of the “shadow,” 1968).

In a review of the early research literature Holmes (1969, 1978) concluded that there was no empirical evidence to support the psychoanalytic model of defensive projection. However, other researchers have criticized Holmes’ conclusion based on an overly stringent definition of defensive projection (Newman, Duff and Baumeister, 1997; Govorun et al., 2006; Schimel et al., 2003). First of all, Holmes assumed that defensive projection required that the individual be unaware of possessing the traits he or she projected onto others. Second, Holmes assumed that projection must effectively reduce an individual’s anxiety stemming from the unwanted trait.

As a consequence of these unnecessarily narrow definitional parameters, the topic of projection was largely ignored by mainstream social psychology until Newman, Duff, and Baumeister (1997) introduced a socio-cognitive re-formulation of the phenomenon. They posited that projection emerges as a byproduct of another defensive process, namely, the active suppression of undesirable traits. They proposed that the thought of having an undesirable trait poses a self-threat which the individual is motivated to alleviate through a process of active suppression. Research on ironic processing effects show that attempts to suppress a thought can lead to its hyper-accessibility (for a review, see Wegner, 1992). Drawing on this research, Newman and colleagues posited that the heightened accessibility of the individual’s undesirable trait resulting from the suppression process influences his or her perception of others. In support
of this model, they found that those with a greater dispositional tendency to use suppression, as well as those induced to suppress a specific trait, both evidenced an increased tendency to project onto another.

A number of more recent studies have provided converging evidence that the motivation to maintain a positive self-image can lead individuals to project their own salient undesirable traits onto others. Perhaps the strongest evidence for this phenomenon comes from research conducted by Schimel and colleagues (2003). In their first experiment, participants were randomly assigned to receive false feedback on a bogus personality test indicating that they either had high or low levels of repressed hostility. Participants then read an ambiguous vignette about a man named Donald whom they were asked to evaluate on an assortment of traits. Three of those traits were related to hostility (hot-tempered, hostile, irate), whereas the remaining traits were unrelated to anger or hostility (e.g., boring, intelligent, likable, narrow-minded, dependable, lazy). As predicted, participants primed to think they possessed high (vs. low) levels of latent hostility rated Donald as being significantly angrier. Importantly, the bogus feedback had no effects on evaluations of Donald on any of the other negative or positive traits.

Govorun and colleagues (2006) extended this work by testing the central role of self-threat. In one study, participants were randomly assigned to write about either an instance in which they failed at an intellectual task, an instance in which an acquaintance failed at an intellectual task, or a typical day for either themselves or an acquaintance. Participants were then asked to indicate whether a variety of stereotypic traits were consistent with a specific group, student athletes. Six of the traits were related to a lack of intelligence (e.g., incompetent, stupid), whereas the remaining traits were not (e.g., motivated, cheerful, dishonest). Compared to participants in the other conditions, participants who recalled their own intellectual failure were
the fastest to endorse traits implying a lack of intelligence to be stereotypic of student athletes. Also, in line with Schimel et al.’s finding that people project a particular salient undesired trait and not global negativity, Goverun et al. found that the salience of personal unintelligence did not affect endorsement of negative stereotypic traits that were unrelated to intelligence (e.g., arrogant). These results suggest that it is the threat of possessing an undesirable trait oneself, and not simply exposure to that trait in another, that elevates one’s tendency to project that trait.

Goverun and colleagues (2006) also examined the characteristics that make someone an acceptable projection target. They hypothesized that people can only project their own undesirable traits onto a person who can be justifiably seen as possessing the trait in question. That is, people should be able to project their own moral inadequacy onto a thief, but not a Good Samaritan. In support of this hypothesis, they found that reminding participants of their own intellectual failings caused them to perceive another person to be less intelligent, but only if that person was a member of a group commonly stereotyped as unintelligent. The researchers explained that the stereotype of low intelligence provided participants with the justification necessary to project their own feelings of intellectual inadequacy onto the target. Importantly, as with their previous study, no effects were found on ratings of stereotypic and non-stereotypic traits unrelated to intelligence. These null effects show that the obtained effects represent the projection of a specific self-threatening trait, and not merely a general tendency to activate stereotypes or derogate stereotyped others in response to a self-threat (Fein & Spencer, 1997).

While the aforementioned research provides compelling evidence for defensive projection, evidence that projection effectively serves a self-protective function is mixed. On the one hand, Schimel et al. (2003) showed that projection can effectively restore a positive personal identity. Specifically, participants told that they possessed an undesirable trait, and subsequently
given the opportunity to project that trait onto another person, showed reduced accessibility of
trait-relevant words (Study 1) and were less likely to ascribe the undesirable trait to the self
(Study 2). On the other hand, a number of other studies have failed to find evidence that
projecting an undesirable self-attribute onto another has any discernible effect on participants’
self-perceptions or felt anxiety (Govorun, Fuegen, & Payne, 2005; Halpern, 1977; Holmes &
Houston, 1971).

In summation, the existing research on projection suggests that although people see their
own salient undesirable qualities in certain others when the possession of these qualities
threatens the self, this process does not necessarily function to alleviate the underlying threat.
Thus, increasing the salience of one’s own immoral actions may produce an increased tendency
to see immorality in another transgressor without necessarily alleviating one’s own perceived
immorality. According to the ancient rituals of purification which form the anthropological
touchstone for the current research, a secondary process is necessary to alleviate one’s moral
contamination, namely, the punishment of the projection target.

Punishment

Punishment has long been described as a means of atoning for sin and alleviating the
stain of immorality. As the Judeo-Christian Bible states: “punishment cleanses away evil”
(Proverbs 20:30, NLT). Indeed, research shows that people view punishment as atoning for
immoral behavior. Nelissen and Zeelenberg (2009) found that when people were induced to feel
that they had harmed others but had no means of making reparations, they administered more
self-punishment by denying themselves a pleasurable reward (Study 1) or giving themselves
point deductions on an ego-relevant task (Study 2). These self-punishment effects were
presumably motivated by participants’ desire to alleviate their increased feelings of guilt.
Similarly, Bastian and colleagues (2011) found that when participants wrote about a time when they had harmed another person, the participants chose to expose themselves to more physical pain by submerging their hands in a bucket of ice water for an extended period of time. The researchers also found that exposure to physical pain under these conditions served to reduce participants’ feelings of guilt, but did not affect other negative emotions (e.g., angry, distressed afraid).

More recently, Inbar, Pizarro, Gilovich and Ariely (2013) examined whether guilt specifically motivates self-punitive behavior. Participants were randomly assigned to describe either a time when they felt “most guilty,” a time when they felt “most sad,” or the last time they went to the grocery store. A manipulation check confirmed that participants in the guilt and sadness condition reported feeling higher levels of guilt and sadness, respectively. Two electrodes were then attached to the participants’ wrists and they were given the opportunity to control the intensity of five self-administered electric shocks. After the shock task, participants reported the extent to which they felt a variety of positive and negative emotions. As predicted, the mean voltage in the guilt condition was significantly higher than the sadness or control condition. Self-reported feelings of guilt prior to the shock task were also correlated with higher voltage intensity. Furthermore, a higher intensity of self-administered shocks was associated with a greater reduction in participants’ self-reported guilt following the shocks. This study provides more evidence that people are motivated to receive punishment as a means of removing aversive feelings of guilt over their own immoral actions.

Because the aforementioned research exclusively assesses self-administered punishment, it is not possible to disentangle the guilt-alleviating effects of being punished from the potential effects of administering punishment. This distinction is important in light of recent research
showing that punishing others can bolster feelings of morality. Adams and Monin (2013) found that participants given the chance to punish a moral transgressor felt more “morally just” compared to those who saw the transgressor go unpunished (Study 1 & Study 2), and those who witnessed the transgressor be punished by a third party (Study 3). This research suggests that the act of punishing others bolsters a punisher’s self-perceived moral identity. Importantly, this effect is conceptually distinct from the idea that punishment serves to atone for the immorality of the transgressor who is punished, regardless of who is administering the punishment.

**Overview of Current Research**

As noted earlier, many pre-modern cultures engaged in purification practices to maintain a positive moral identity, and while these practices are diverse in appearance, they are typically based on three beliefs: a) guilt over one’s immoral actions is a tangible contaminant; b) one’s own moral contamination can be placed onto others; and c) immorality is eliminated through the reception of punishment (Douglas, 1995). Although such beliefs may seem unusual today, the three lines of research just reviewed suggest these beliefs continue in the modern era to shape people’s moral judgments and perceptions. This raises the possibility that, even among modern individuals, witnessing another person being punished for a moral transgression may serve to “cleanse” the self of immorality. More specifically, we hypothesized that when people feel tainted by their immoral actions, they are motivated to see their own moral contamination in another transgressor; furthermore, the punishment of that transgressor serves to restore their own moral and physical purity.

Two experiments tested variants on this broad hypothesis. In Study 1, we predicted that focusing participants on their immoral actions would cause them to report feeling physically contaminated, and this effect would be attenuated if they were presented with another moral
transgressor who was punished (but not an unpunished transgressor or a non-transgressor). In Study 2, we predicted that participants whose own immoral actions were made salient would perceive another target transgressor as more physically contaminated, and that this effect would occur indirectly through participants’ own increased feelings of personal guilt. We also predicted that participants primed with their own immoral actions would report decreased perceptions of personal morality and an increased desire to engage in compensatory moral behavior, but these effects would be attenuated if the target transgressor was punished by a third party.

**Study 1**

Zhong and Liljenquist (2006) found that individuals who thought about their own immoral actions showed an increased desire to physically cleanse themselves. Based on our claim that the punishment of another moral transgressor “cleanses” the self of moral contamination, we hypothesized that increasing the salience of personal immoral actions would cause participants to feel physically dirtier, but not if they were additionally exposed to another transgressor who is punished. Because of the unique, morally cleansing effect of punishment, we did not expect this effect to occur when one was exposed to a non-punished transgressor, even when the extent and severity of the target’s wrongdoing was the same as that of the punished transgressor. By presenting equivalent transgressors and manipulating punishment, Study 1 sought to rule out downward social comparison as an alternative explanation for the predicted effects. If these effects were simply due to downward social comparison, we would expect participants to feel less guilty about their own immoral actions after observing another person’s immoral action, regardless of whether that person’s immoral action was punished or not.
Method

Participants and design. Participants were 153 American adults (76 female) ranging in age from 18 to 82 years ($M = 32.87, SD = 13.40$) who received $0.35 for completing an online survey made accessible through Amazon’s Mechanical Turk (Mturbk) service. The experiment was described as three unrelated studies investigating different aspects of memory, judgment, and personality. Participants were randomly assigned to conditions in a $2 \times 3$ between-subjects factorial design, with self-reported feelings of personal physical cleanliness serving as the dependent variable of primary interest.

Personal immorality salience manipulation. As part of a supposed study on memory, participants responded to a prompt that asked them to recall and write about a specific time in their life. Participants in the immorality salient condition responded to a writing prompt based on Zhong and Liljenquist’s (2006) materials (see Appendix A). Specifically, they were asked to describe a time when they acted “in an unethical way.” Participants in the immorality-not-salient condition responded to a parallel prompt that asked them to describe a time when they were “very bored,” an aversive topic intended to control for the general negativity of the immorality salience induction. All participants were instructed to write three to five sentences describing the situation, how they acted, and how it made them feel. In an attempt to control participants’ temporal distance from the recalled event, they were instructed to think of an event that occurred at least one year ago, but after their childhood.

Target manipulation. As part of an ostensibly unrelated second study, participants were then presented with a report from what was purported to be a university student misconduct
hearing. The first half of the report, which was the same for all participants, provided a case description of an incident in which a student was accused of stealing money from a charity donation box on campus (see Appendix B). The second half of the report which described the final judgment and disciplinary action taken by the committee differed according to condition (see Appendix C). Participants in the non-transgressor condition read that the student was found not guilty and that all the money was accounted for. Participants in the non-punished transgressor condition read that the student was found guilty of the offense, but no punitive action could be taken because the individual was no longer a student at the university. Participants in the punished transgressor condition read that the student was found guilty and was expelled for the offense. All participants then rated the accused student on three negative traits (i.e., stupid, lazy, close-minded) along a 6-point scale (1 = not at all, 6 = very much; \( \alpha = .67 \)). These traits were assessed to ensure that participants perceived the punished and non-punished targets as equally negative. Because these negative traits were not related to the target’s moral standing, we did not expect them to be affected by the personal immorality salience prime.

**Personal physical cleanliness measure.** Next, as part of another ostensibly unrelated third study, participants were given a questionnaire instructing them to indicate their agreement with four statements about their current feelings of personal physical cleanliness: “I feel exceptionally clean”; “I feel very sanitary”; “I feel filthy” (reverse scored); “I feel polluted” (reverse scored). Responses were made on a 6-point scale (1 = strongly disagree, 6 = strongly agree) and were averaged (after reverse-scoring the latter two) to form composite scores (\( M_{\text{grand}} = 4.42, SD = .81; \alpha = .76 \)).
Results

Target ratings. To test our assumption that the punishment of a transgresser did not influence general negative attitudes toward the transgresser, we submitted negative trait ratings to a 2 (immorality salient vs. immorality-not-salient) × 3 (non-transgressor vs. non-punished transgressor vs. punished transgressor) ANOVA. As expected, we observed only a main effect of target type, $F(2, 148) = 18.26, p < .001$, such that the negative ratings for the punished transgressor ($M = 4.52, SD = .84$) and non-punished transgressors ($M = 4.37, SD = 1.03$) were higher than the negative trait ratings for the non-transgressor ($M = 3.29, SD = 1.31; ps < .001$). Negative ratings of the two transgressor conditions did not differ ($p > .41$).

Personal physical cleanliness. To test our primary contamination and cleansing hypotheses, we submitted ratings of personal physical cleanliness to the same personal immorality salience × target ANOVA. This analysis revealed a main effect of immorality salience, $F(1, 148) = 4.68, p = .03$, which was qualified by the predicted two-way interaction, $F(2, 148) = 4.98, p = .01, \eta^2_p = .06$ (see Figure 1 for the pattern of means).

In line with previous research (e.g., Zhong & Lilenquist, 2006), pairwise comparisons (Fisher’s LSD) revealed that in the non-transgressor condition, participants primed with their own immoral actions reported feeling significantly less physically clean ($M = 4.07, SD = .73$) compared to participants whose personal immorality was not made salient ($M = 4.60, SD = .67; F(1, 148) = 5.63, p = .02$). Similarly, in the non-punished transgressor condition, participants primed with their own immoral actions reported feeling significantly less physically clean ($M = 4.06, SD = 1.08$) compared to participants whose personal immorality was not made salient ($M = 4.67, SD = .69; F(1, 148) = 7.60, p = .01$).
In contrast, in the punished transgressor condition, participants’ feelings of personal physical cleanliness did not differ according to whether their own personal immorality was made salient ($M = 4.62$, $SD = .50$) or not ($M = 4.32$, $SD = .93$; $F(1, 148) = 1.72, p = .19$). Also supporting predictions, in the immorality salient condition, participants exposed to a punished transgressor reported feeling significantly cleaner than those exposed to either a non-transgressor or non-punished transgressor ($F > 4.49$, $ps < .04$). No effect of target type on felt physical cleanliness was found in the immorality-not-salient condition ($F < 2.70$, $ps > .10$).

**Study 1 Discussion**

Supporting predictions, the salience of personal immoral actions led participants to report feeling physically dirtier, but this effect was eliminated if participants were additionally presented with another transgressor who was punished, but not a non-transgressor or a transgressor who was not punished. These results are consistent with Zhong and Liljenquist’s (2006) research suggesting that people can be physically soiled by their own immoral actions, but they go further to support our claim that the punishment of another moral transgressor by a third party can effectively eliminate the feelings of personal physical contamination evoked by one’s own salient personal transgressions.¹

Study 1 also helps to rule out social comparison processes as an alternative mechanism behind our observed effects. If these effects were simply due to comparing the self with someone who is worse off, we would have expected that exposure to the same transgressor would similarly influence participants’ self-perceptions, regardless of whether or not that transgressor was ultimately punished. Yet the results of Study 1 show that punishment, and not simply downward social comparison, is necessary to cleanse the self of immoral actions. Still, the social comparison alternative has not been definitively ruled out. It is possible that despite equivalent
descriptions and non-moral negativity ratings of the transgressor across both the punished and non-punished target conditions, participants judged the punished transgressor as dirtier or more evil than the non-punished transgressor. This leaves open the possibility that our effects are due to downward social comparison on these dimensions.

**Supplemental Study**

To rule out this possibility, we randomly assigned another group of 65 participants to one of the three target conditions used in Study 1 and asked them to rate the extent to which the target could be characterized as dirty and as evil (1 = not at all, 6 = very much). Dirty and evil scores were submitted to separate ANOVAs by target condition (see Table 1 for means and standard deviations). Results showed that the non-punished transgressor target was rated as dirtier and more evil than both the punished transgressor target (ps < .04) and the non-transgressor target (ps < .02). Ratings of the punished transgressor target and non-transgressor target did not differ on either characteristic (ps > .70). Additionally, replicating Study 1, an ANOVA on composite non-moral negative trait ratings (stupid, rude, conceited; \( \alpha = .77 \)) found that the punished transgressor target (\( M = 4.14, SD = .87 \)) and non-punished transgressor target (\( M = 4.29, SD = 1.15 \)) were given greater negative trait scores than the non-transgressor target (\( M = 2.76, SD = 1.16; ps < .02 \)), though ratings of the two transgressor targets did not differ (\( p = .34 \)).

Although there were significant differences in the perceived morality and physical cleanliness of punished and non-punished transgressor targets, the pattern of these differences is inconsistent with a downward comparison account of the results obtained in Study 1. In fact, based on these supplemental findings, a social comparison perspective would predict that participants exposed to a non-punished transgressor target should feel comparably cleaner than
those exposed to a non-transgressor or punished transgressor target. In contrast, Study 1 found that when primed with their own immoral actions, participants felt less clean when exposed to a non-punished transgressor and non-transgressor compared to when exposed to a punished transgressor. This provides further evidence that Study 1’s effects were not the result of downward comparison processes.

The results of this supplemental study also provide empirical support for our claim that punishment is perceived to expiate the moral transgressions of the punished target, restoring the target’s physical and moral purity. This moral cleansing function of punishment is critical to our claim that the punishment of another can serve to cleanse the self when contamination elicited by one’s own immoral actions is projected onto the punished target. While the results of Study 1 are consistent with this explanation, because the primary study never assessed the perceived cleanliness of the target, this study does not provide direct evidence that participants primed with their own immoral actions actually project their own felt contamination onto the target transgressor. We test this directly in Study 2.

Study 2

Study 2 had three primary goals. The first goal was to directly test whether people project their own salient immorality, which is experienced as a physical contamination (Study 1), onto another transgressor. To do this, we had all participants evaluate the same moral transgressor in the absence of any punitive information following the personal immorality salience manipulation. We hypothesized that participants who contemplated their own immoral actions would perceive the transgressor as physically dirtier as the result of their own increased feelings of personal guilt. Based on previous research on projection (Govorun et al., 2006; Schimel et al., 2003), we also hypothesized that the personal immorality prime would not influence ratings of
the target transgressor on negative traits unrelated to their moral standing. Null results on non-moral negative evaluations of the transgressor would serve to rule out the possibility that rating another transgressor as physically dirty is simply an attempt to derogate the target. This is an important distinction since previous research has shown that derogating stigmatized others can bolster threatened feelings of self-esteem (Fein & Spencer, 1997).

Study 1 found that exposure to a punished transgressor effectively restored participants’ perceived physical cleanliness following a salient personal transgression. The second goal of Study 2 was to test whether this effect could be replicated for participants’ perceived personal morality. We hypothesized that participants reminded of their own immoral behavior would perceive themselves as less moral unless they learned that the transgressor on whom they had projected their own moral contamination was punished. As in Study 1, punishment of the transgressor was always administered by a third party. This was done to ensure that any effects of punishment could not be explained by the moralizing effects of administering punishment to others (Adams & Monin, 2013).

The third goal of Study 2 was to investigate an important potential downstream consequence of the hypothesized moral cleansing process. Previous research has shown that people attempt to alleviate felt guilt over their own misdeeds by engaging in compensatory acts of prosocial behavior (Carlsmith & Gross, 1969; Darlington & Macker, 1966; Regan, Williams, & Sparling, 1972). As discussed in the Introduction, Zhong and Lilenquist (2006) found that the act of physical cleansing ameliorated participants desire to engage in compensatory prosocial behavior in response to a reminder of their own immoral actions. Insofar as the punishment of another transgressor serves to morally “cleanse” the self, we hypothesized that participants primed with their own immoral behavior would report an increased desire to engage in prosocial
behavior (i.e., volunteer in a local blood drive), but this effect would be eliminated if they were additionally presented with a punished transgressor.

Method

Participants and design. Participants were 87 undergraduates (42 women) from a Midwestern university ranging in age from 18 to 34 years ($M = 22.21$, $SD = 3.57$), who participated in partial fulfillment of a course requirement. As with Study 1, the experiment was presented as three unrelated studies investigating different aspects of memory, judgment, and personality. In private cubicles, participants completed the first two of the purported three studies on computers while the materials for the third purported study were completed on paper and pencil surveys. Participants were randomly assigned to conditions in a 2 (personal immorality salience: immorality salient vs. immorality-not-salient) $\times$ 2 (target type: punished transgressor vs. non-punished transgressor) between-subjects factorial design.

To test for projection, self-reported feelings of personal guilt were assessed as the hypothesized mediator with ratings of the perceived physical dirtiness of the target transgressor as the primary dependent variable. To test the moral cleansing hypothesis, the primary dependent variables were participants’ rating of their personal moral character and their reported interest in engaging in prosocial behavior (i.e., joining a local blood drive).

Personal immorality salience manipulation. As part of a supposed study on memory and emotions, participants first responded to the same immorality salience manipulation used in Study 1. Participants randomly assigned to the immorality salient condition were asked to describe a time when they acted “in an unethical way,” whereas participants in the immorality-not-salient condition were asked them to describe a time when they were “very bored.”
**Guilt measure.** Following this manipulation, participants were asked to indicate the extent to which they currently felt guilty using a 5-point scale (1 = *very slightly or not at all*, 5 = *extremely*). In addition to completing this single item measure of personal guilt ($M_{\text{grand}} = 1.54$, $SD = .86$), participants also rated the extent to which they felt 19 other emotions included in Watson, Clark and Tellegen’s (1988) Positive and Negative Affect Schedule (PANAS). Ten of these items were positive emotion words (interested, alert, excited, inspired, strong, determined, attentive, enthusiastic, active, proud) and the remaining nine were negative emotion words (angry, distressed, upset, nervous, ashamed, scared, hostile, jittery, afraid). Separate composite scores were computed for the positive emotion subscale ($M_{\text{grand}} = 2.51$, $SD = .87$; $\alpha = .91$) and the negative emotion subscale excluding guilt ($M_{\text{grand}} = 1.54$, $SD = .55$; $\alpha = .84$).

**Target manipulation.** Next, participants were told that they would take part in a purportedly unrelated study on memory and judgment that involved reading short news articles. Participants were first presented with an experimenter-fabricated news story about a hit-and-run car accident in which a 25 year old man drove through a red light and hit a pedestrian before fleeing the scene (see Appendix D). The article concluded by saying that the driver had been identified but not yet apprehended. In addition to detailing the crime, the article also included a photograph of the driver who represented the target transgressor in the current study.²

After reading the article, participants were presented with an ostensible memory test that asked them to recall information about the article they had read. As part of this test participants were asked to, “think back to the photograph of the perpetrator (that is, the driver who hit someone)” and rate their agreement with various statements about his appearance. Two of the items concerned the transgressor’s physical dirtiness: “The perpetrator appeared to be dirty”; “The perpetrator appeared to have poor hygiene.” Responses were made on a 6-point scale (1 =
strongly disagree, 6 = strongly agree) and were averaged to form composite scores ($M_{\text{grand}} = 3.39, SD = 1.33; \alpha = .87$).

To assess participants’ general negative evaluation of the transgressor, we also instructed participants to respond to three items rating the extent to which the transgressor was “incompetent,” “stupid,” and “clever” (reverse scored). Responses were made on a 6-point scale (1 = not at all, 6 = very much) and were averaged to form composite negative trait scores ($M_{\text{grand}} = 5.15, SD = 1.06; \alpha = .84$). These traits were assessed to ensure that the personal immorality prime was not simply increasing participants’ general tendency to derogate the transgressor. As in Study 1, given the specificity of the projection phenomenon, we did not expect ratings of these non-moral negative traits to be affected by the personal immorality salience manipulation.

Next, all participants were randomly assigned to be presented with one or two possible experimenter-fabricated news articles supposedly written 6 months after the first article that provided updated information about the aforementioned hit-and-run accident (see Appendix E). This article comprised our target type manipulation. Participants randomly assigned to the punished transgressor condition read that the driver had been apprehended, convicted of various charges and was currently serving a three year prison sentence. Participants randomly assigned to the non-punished transgressor condition read that the driver had not been apprehended, but if caught, could face various charges and a three year prison sentence.

**Personal morality measure.** Participants were then presented with paper survey packets purported to be part of another unrelated personality study. Included in this packet was a modified version of a questionnaire constructed by Zhong and colleagues (2010) which asked participants to rank themselves in comparison to other undergraduates at their university on eight different dimensions (sense of humor, intelligence, moral character, creativity, physical...
attractiveness, physical fitness, social sensitivity, leadership). Participants responded to each item by indicating the percentile that described their position relative to others, from 0 (worse than all others) to 100 (better than all others). In line with Zhong et al. (2010), we used participants’ response to the moral character item as a measure of their perceived personal morality ($M_{\text{grand}} = 80.24, SD = 10.01$).

**Prosocial behavior intention measure.** At the end of the study, participants read about a community outreach program designed to engage college students in local volunteering opportunities. Participants were asked to indicate whether or not they would be interested in participating in a local blood donation drive by marking a Yes or No response box. Their response to this item provided a categorical assessment of participant’s willingness to engage in prosocial behavior.

**Results**

**Guilt.** We first conducted a univariate ANOVA on self-reported feelings of guilt. We did not include target type as an independent variable in this analysis because the target manipulation came after assessing participants’ guilt. However, a Levene’s test for heterogeneity of variance indicated that responses to our measure of personal guilt violated the homogeneity of variance assumption, $F(1, 85) = 21.88, p < .001$. Accordingly, we conducted Welch’s alternative ANOVA procedure on personal guilt scores, which Tomarken and Serlin (1986) identify as the optimal procedure when the homogeneity of variance assumption is violated. As predicted, this analysis revealed that participants whose immorality was made salient reported feeling significantly more guilt ($M = 1.88, SD = 1.00$) compared to participants whose immorality was not made salient ($M = 1.20, SD = .51$), $F(1, 61.93) = 15.70, p < .001$. Importantly, no effect of personal immorality salience was found on either subscale of the PANAS (positive subscale: $F < 1, p > .88$; negative
subscales without guilt: $F < 1, p > .49)$. Furthermore, the primary effects of immorality salience on guilt remained significant when simultaneously controlling for the positive and negative subscale scores ($p < .001$).

**Physical dirtiness of target transgressor.** To test the prediction that participants reminded of their own immorality will perceive another transgressor as more physically contaminated, we conducted a univariate ANOVA on participants’ ratings of the physical dirtiness of the target transgressor. Once again, we did not include target type in this analysis because this manipulation came after participants’ made their evaluations of the target. As predicted, this analysis revealed that participants in the immorality salient condition rated the transgressor as physically dirtier ($M = 3.77, SD = 1.19$) than participants in the immorality-not-salient condition ($M = 3.02, SD = 1.35; F(1, 85) = 7.37, p = .01, \eta^2_p = .08$). To test whether this effect reflects a tendency to globally derogate the target, we submitted general negative trait ratings of the transgressor to the same ANOVA. As expected, we observed no significant effect of immorality salience on non-moral negative ratings of the transgressor ($F < 1, p > .53$).

**Mediation of personal immorality salience on perceived physical dirtiness of transgressor by guilt.** We conducted an indirect effect analysis to test our mediation hypothesis that the increased perceived physical dirtiness of the transgressor in the personal immorality salience condition occurs through participants’ increased feelings of personal moral contamination.

Using Preacher and Hayes’ (2008) bootstrapping procedure, we regressed perceived dirtiness of target transgressor onto the effect of personal immorality salience (coded: immorality salient = 1; immorality-not-salient = 0) with guilt scores entered as the proposed mediator. Five-thousand bootstrap resamples were performed. The 95% confidence interval obtained for the
indirect effects of immorality salience on perceived transgressor dirtiness scores through guilt did not contain zero (.07, .50). These results are consistent with the mediation hypothesis that the increase in participants’ ratings of the transgressor’s physical dirtiness in response to participants’ own salient transgressions occurred indirectly through a corresponding increase in feelings of personal guilt (see Figure 2 for a graphical depiction of the model).

**Personal morality.** We then tested the prediction that the punishment of a transgressor imbued with one’s own felt contamination can restore one’s positive moral identity. We submitted participants’ self-reported ratings of their own morality to a 2 (personal immorality salience) × 2 (target type) ANOVA. This analysis yielded significant main effects for both immorality salience ($F(1, 83) = 10.13, \ p = .002$) and target type ($F(1, 83) = 9.50, \ p = .003$), which were qualified by the predicted two-way interaction, $F(1, 83) = 16.85, \ p = .01, \ \eta_p^2 = .16$ (see Figure 3 for the pattern of means). Importantly, this interaction remained significant when controlling for participants’ comparative rankings of their standing on morally-irrelevant but positive traits (e.g., creativity; $p = .01$).

Pairwise comparisons revealed that when exposed to a non-punished transgressor, participants primed with their own immoral actions reported lower ratings of their own relative moral character ($M = 70.71, \ SD = 8.98$) compared to those whose personal immorality was not made salient ($M = 84.05, \ SD = 8.61; \ F(1, 83) = 25.68, \ p < .001$). In contrast, when exposed to a punished transgressor, participants’ ratings of their own moral character did not differ according to whether their own personal immorality was made salient ($M = 83.86, \ SD = 8.85$) or not ($M = 82.17, \ SD = 7.66; \ F(1, 83) < 1.00, \ p = .51$). Also, consistent with the primary predictions, for participants whose personal immorality was made salient, exposure to a punished (vs. non-punished) transgressor increased their ratings of their own moral character ($F(1, 83) = 25.56, \ p < $
No effect of punishment on personal morality ratings was found in the immorality-not-salient condition ($F(1, 83) < 1.00, p = .47$).

**Prosocial behavior intentions.** See Table 2 for distribution of responses to the opportunity to engage in prosocial behavior (local volunteering). We regressed these responses onto personal immorality salience (immorality salient vs. immorality-not-salient), target type (punished transgressor vs. non-punished transgressor), and their interaction using a logistic regression analysis. This analysis yielded a significant two-way interaction, $B = -2.48, SE = 1.02, Wald = 5.94, p = .02$ (see Figure 4 the pattern of odds ratios).

Follow-up analyses revealed that when exposed to a non-punished transgressor, the odds of reporting an interest in joining a blood donation drive (saying yes compared to no), when participants’ personal immorality was made salient were 3.52 times the odds of expressing an interest when their immorality was not made salient, $B = -1.26, SE = .67, Wald = 3.49, p = .06$. In contrast, when exposed to a punished transgressor, differences in the odds ratios for those in the immorality salient and immorality-not-salient conditions did not statistically differ ($B = 1.21, SE = .76, Wald = 2.56, p = .11$).

Also consistent with the primary predictions, for participants whose personal immorality was made salient, the odds of expressing interest in joining a blood donation drive when exposed to a non-punished transgressor were 6.97 times the odds of expressing interest in joining a blood donation drive when exposed to a punished transgressor, $B = -1.94, SE = .76, Wald = 6.53, p = .01$. In contrast, for participants whose personal immorality was not made salient, the odds ratios did not differ between those exposed to a punished or non-punished transgressor ($B = .54, SE = .67, Wald = 0.63, p = .43$).
Study 2 Discussion

Supporting our first primary prediction, the salience of personal immoral actions led participants to feel increased personal guilt, which in turn led them to perceive another moral transgressor as physically dirtier. This finding shows that guilt over one’s own immoral actions is grounded in the concrete experience of physical contamination. Additionally, it shows that participants are motivated to project this contamination onto another transgressor. The null effects of immorality salience on both self-reported negative emotions (excluding guilt) and non-moral negative ratings of the target transgressor rule out the possibility that the observed effects simply reflect variation in negative affect or global derogation of a target in response to a self-relevant threat.

The obtained effects on participants’ perceived personal morality were consistent with the effects obtained in Study 1 on participants’ perceived physical cleanliness. Supporting our second prediction, the salience of personal immoral actions led participants to rate themselves as less moral compared to peers, unless they read that the transgressor who served as a projection target had been punished for his crimes. These results provide evidence that the punishment of another transgressor can function to alleviate one’s own felt immorality. Distinguishing this phenomenon from more general self-esteem maintenance or enhancement effects, the primary findings remained significant when controlling for participants’ self-evaluations in non-moral domains. Additionally, the fact that participants were explicitly asked to compare their own moral character with that of their fellow students makes it unlikely that the obtained results are the result of a downward social comparison with the transgressor.

Supporting our third prediction, the salience of personal immoral actions motivated an increased willingness to express interest in volunteering in a local blood drive, unless the other
transgressor was punished. These results mirror Zhong and Liljenquist’s (2006) finding that physical cleansing alleviated people’s willingness to volunteer to help another student in need. This suggests that like physical cleansing, the punishment of a transgressor perceived to possess one’s own felt contamination also serves a moral cleansing function for the self. Furthermore, although we did not directly measure participants’ actual prosocial behavior, the results of Study 2 suggest that the punishment of other transgressors can, under certain conditions, have a deleterious effect on people’s own expressions of prosociality.

**General Discussion**

The two studies supported the broad hypothesis that when people feel tainted by their immoral actions, they are motivated to see their moral contamination in another transgressor; furthermore, the punishment of that transgressor serves to restore their own moral and physical purity. Study 1 found that when participants contemplated their own immoral actions they felt less physically clean, unless they were exposed to another wrongdoer who was punished for his transgression. Study 2 found that participants who contemplated their own immoral actions perceived another wrongdoer as physically dirtier via increased feelings of personal guilt, and felt less moral and more willing to engage in moral restoration behavior unless they learned that the contaminated transgressor had been punished. Importantly, the increased feelings of physical and moral contamination elicited by personal immorality salience were not mitigated by exposure to a moral transgressor who was not punished. These studies suggest that personal immoral actions induce feelings of contamination that individuals are motivated to project onto a target whose punishment is perceived to remove their own contamination.

Social comparison theory (Festinger, 1954) claims that people can restore or maintain a positive self-image by comparing themselves with others who are worse off than themselves
(i.e., making a downward social comparison). Social comparison research shows that people are motivated to engage in a downward social comparison when their self-image is threatened (Wills, 1981). This raises the possibility that people threatened with their own immoral behavior in the current studies may have perceived themselves as cleaner or less immoral by making a downward social comparison with the target transgressor who was perceived to be dirtier and more immoral than they were.

We attempted to rule out this possibility by exposing participants to equivalent transgressors and only manipulating whether or not the transgressor was punished. Although perceptions of the transgressor were not directly assessed in Study 1, a supplemental study found that participants rated the punished transgressor as cleaner and less evil than the non-punished transgressor. Thus, if participants were engaged in a social comparison with the target transgressor, they should have felt comparably cleaner when exposed to the non-punished (vs. punished) transgressor. However, Study 1 found that participants in the personal immorality salience condition reported feeling physically cleaner when exposed to a punished (vs. non-punished) transgressor. Study 2 attempted to ensure that participants’ did not engage in downward social comparison with the target transgressor by having participants rate their moral character in comparison with a different target, their student peers.

To further establish discriminant validity, we took steps to distinguish the phenomenon of projection from that of defensive derogation of others. Previous research by Fein and Spencer (1997) found that participants who received negative feedback on a self-relevant performance task were more likely to derogate a stereotyped target, in what they described as a kind of defensive prejudice. They also found that derogating a target in response to negative self-relevant feedback served to increased participants’ self-esteem. This raises the possibility that
participants’ tendency to rate a target transgressor as physically dirty when reminded of their own immoral actions in the current studies may have been part of an effort to derogate the target in order to bolster their own flagging self-esteem. If this were the case, we would expect personal immorality salience to increase general negative evaluations of the target transgressor as well as positive self-evaluations. In contrast to this prediction, Study 2 found that personal immorality salience had no effect on participants’ global negative ratings of the target transgressor and no effect on participants’ non-moral self-ratings. In this way, the current research meets an empirical criterion established in prior research (Govorun et al., 2006; Schimel et al., 2003) for distinguishing defensive projection from self-threat-induced derogation: the increased salience of an undesirable trait in the self, influenced evaluations of a target individual specifically with regard to that trait, and not with regard to other negative traits as would be the case in global derogation.

Taken together, the results of the present studies provide evidence for a moral cleansing mechanism that guides moral attitudes and interpersonal perceptions in a similar manner today as it appears to have done in pre-modern cultures across the world. While research on metaphor, projection, and punishment have shown that the core assumptions of pre-modern cleansing rituals remain relevant, our data show that motivations to place one’s own immorality onto others and see them punished may continue to play an important role in the pursuit and maintenance of moral cleanliness. In addition to providing initial evidence for a practically important self-serving process behind moral reasoning, the current research broadens the theoretical and empirical scope of the three research areas it builds upon, as described in the next sections.

**Advancing Conceptual Metaphor Theory and Research**
Research investigating CMT has provided a wealth of evidence indicating that people draw on their knowledge of a familiar, concrete concept to understand superficially unrelated abstract concepts (see Landau, Meier, & Keefer, 2010). Studies have shown that manipulating experience with a concrete concept (e.g., inducing sensory and motor states) produces metaphor-consistent effects on perceptions related to an abstract concept (e.g., Lee & Schwarz, 2011; Schnall et al., 2008; Zhong et al., 2010). Likewise, manipulating representations of an abstract concept can produce metaphor-consistent effects on perceptions and sensations related to a concrete concept (Lee & Schwarz, 2010; Zhong & Liljenquist, 2006). Lee and Schwarz (2012) claim that a juxtaposition of these separate lines of work reveals a bidirectional relationship between the abstract and concrete which emerges through the co-activation of abstract and metaphorically associated concrete concepts. However, since investigations of concrete-to-abstract effects and abstract-to-concrete effects have generally been studied in isolation, there is a dearth of evidence directly assessing this co-activation hypothesis.

The present studies included measures assessing both the abstract and associated metaphoric concept. In support of Lee and Schwarz’s (2012) co-activation hypothesis, we found that the manipulation of the abstract concept produces parallel effects on both morality-relevant and cleanliness-relevant outcome measures. Specifically, Study 2 found that a personal immorality prime elicited both feelings of guilt and perceptions of physical contamination. Additionally, Study 1 and 2 found identical interaction effects on participants’ perceptions of their own physical cleanliness and personal morality. In support of a bidirectional relationship, we found that the manipulation of an abstract concept (immorality salience) produced a complementary sensorimotor experience (decreased physical cleanliness), which was influenced by a subsequent manipulation of the abstract concept (expiating punishment).
The present studies directly contribute to research on the specific metaphorical link between morality and cleanliness by showing that, in addition to cleansing oneself, the punishment of others viewed as immoral represents a strategy for metaphorically “cleansing” the self of moral wrongdoing. These studies also provide a novel interpretation of previous research showing that exposure to unclean physical conditions can motivate people to make harsher moral judgments. Schnall and colleagues (2008) proposed that this effect represents people’s tendency to misattribute physical disgust induced by an unclean environment (Study 2) or a foul smell (Study 1) as moral disgust caused by another transgressor. However, the present research suggests that the harsh judgments of a moral transgressor may represent a motivated desire to purge one’s own felt contamination by projecting it onto a punishable transgressor. In contrast, an attempt to explain the results of the current studies as a tendency to misattribute the cause of disgust over one’s own immoral actions to the physical contamination of a target transgressor does not explain why the punishment of that transgressor serves to morally and physically “cleanse” the self.

Advancing Projection Theory and Research

After a long period of neglect, Newman and colleagues’ (1997) reintroduced the phenomenon of projection as an acceptable topic for empirical inquiry. A number of recent studies investigating projection have found evidence that the desire to maintain a positive self-image can motivate individuals to see in others the undesirable traits they fear they themselves possess. However, support for the defensive function of projection has been mixed. Some studies have found evidence that projection can reduce the projector’s perceived possession of the undesired trait (Schimel et al., 2003); other studies have not (Govorun et al., 2005; Halpern, 1977; Holmes & Houston, 1971). In light of these inconsistent findings, the very existence of
projection as envisaged by generations of psychoanalytic and existentially-oriented theorists as a defensive mechanism seems to have fallen in doubt.

The present research sheds new light on this controversy by demonstrating that attributing one’s own undesirable self-attributes onto a target other is necessary but not sufficient to restore one’s positive moral identity. In addition, the person needs to observe the punishment of the projection target. This is consistent with Freud’s (1915/1957) observation that projection is often characterized by a desire to aggress against the projection target. The present studies add to Freud’s theorizing by showing that aggression towards the projection target can be realized in punishment meted out by a third-party which serves an essential moral cleansing function.

This enhanced conception of defensive projection which includes both the ascription of one’s immorality onto a projection target and the punishment of that target represents a significant contribution to the existing theory and future research on projection. For one, this proposed two-part process provides a useful way to distinguish projection from two mechanisms that are closely related on the surface: social comparison processes and self-threat-induced derogation. This enhanced conception of defensive projection also highlights the need to consider the role of post-projection processes such as the punishment of a projection target. Ultimately, the present research sets a new agenda for taking a fresh empirical look at projection as a unique and practically important process that guides moral attitudes and interpersonal perception.

**Advancing Moral Punishment Theory and Research**

Recent studies have found evidence that punishment is perceived to atone for one’s immoral behavior (Bastian et al., 2011; Inbar et al., 2013; Nelissen & Zeelenberg, 2009). This research shows that when an individual’s own immoral behavior is made salient, they are
motivated to self-punish as a means of reducing increased feelings of guilt. However, as
discussed in the Introduction, because this literature focuses exclusively on self-punishment,
these studies confound the expiating effect of being punished with the act of administering
punishment. This is particularly problematic in light of recent research showing that the
opportunity to administer punishment to another transgressor bolsters one’s own perceived moral
identity (Adams & Monin, 2013).

To our knowledge, the present research provides the strongest evidence to date that
punishment is perceived to alleviate the immorality of the punished target. These studies
accomplished this by exposing participants to a target transgressor who was punished by a third
party. Results of the supplemental study showed that participants perceived a punished
transgressor to be significantly less evil than an equivalent non-punished transgressor, and on par
with a non-transgressor. The punishment of the transgressor not only reduced his perceived
immorality, but appeared to effectively absolve the transgressor of sin. Study 2 also found
indirect evidence that punishment removes the punished transgressor’s moral taint by showing
that the third-party punishment of the target transgressor on whom participants had projected
their own moral contamination served to vicariously restore participants’ perceived moral
identity. At the same time, consistent with Adams and Monin’s (2013) findings, in the absence of
personal immorality salience, reading that a target transgressor had been punished by a third
party had no effect on participants’ perceived morality.

Limitations and Future Directions

Although the current studies provide strong initial evidence for the hypothesized moral
cleansing function of punishing moral transgressors, there are a few notable limitations that
should be addressed by future research. We propose that the punishment of a target transgressor
metaphorically cleanses the self of moral contamination because the transgresser is seen as possessing one’s own contamination. The present studies manipulated the punishment of a target transgressor (Studies 1 & 2) and measured participants’ tendency to project their own felt contamination onto the transgresser (Study 2). As a consequence, these studies illustrate that people project their own felt contamination onto another transgresser and are cleansed by the transgresser’s punishment. However, because we measured, rather than manipulated, participants’ tendency to project onto the transgresser, these studies did not directly test the hypothesized causal role of projection in the moral cleansing process. It remains possible that the punishment of a target other may bolster one’s perceived personal morality and physical cleanliness in the absence of projection. Perhaps simply reading about another transgresser’s punishment in the current studies increased the accessibility of morality- and cleanliness-relevant thoughts which in turn influence participants’ self-perceptions.

This explanation seems unlikely, given that exposure to a punished transgresser did not affect participants’ physical cleanliness or personal morality ratings in the absence of a personal immorality salience induction. However, the hypothesized role of projection could be directly assessed by manipulating whether or not participants are able to project their own felt contamination onto a another target. Govorun and colleagues (2006) found that people are only able to project onto a target who can be justifiably seen to possess the self-attribute people wish to deny in themselves. This implies that participants should be unable to project their own felt moral contamination onto a target perceived to be innocent of wrongdoing. Thus, we would predict that participants reminded of their own immoral behavior would feel morally and physically cleaner when exposed to the punishment of a target transgresser (in whom they can see their own moral taint), but not when exposed to the punishment of a patently innocent target.
The moral cleansing phenomenon evidenced in the present studies is explained as a motivated mechanism for eliminating a perceived threat to one’s own moral identity. In support of this conceptualization, the primary cleansing effects occurred in response to a personal moral value threat. However, our ability to draw firm conclusions about the motivational roots of this process is limited by various methodological factors. For instance, these studies did not include a comparison condition where participants were primed with immorality salience independent of their own behavior. Therefore, it remains possible that simply priming the concept of immorality in the absence of a self-threat would produce the same effects. Importantly, Govorun and colleagues’ (2006) research on defensive projection disputes this claim. Specifically, these researchers found that whereas participants who wrote about their own intellectual failure projected unintelligence onto an applicable target, writing about the intellectual failure of an acquaintance produced no discernible effect.

Additionally, manipulating the third-party punishment of a transgressor allowed us to isolate the unique moral cleansing effect of punishment but did not allow us to assess whether participants were actually motivated to see the transgressor punished. Investigating this possibility would require giving participants the opportunity to voice support for another transgressor’s punishment. Insofar as participants are motivated to see a moral transgressor punished to eliminate their own felt moral contamination, we would expect a personal immorality salience induction to drive increased support for punishing a target transgressor via increased feelings of personal guilt. We would also expect that having an alternative means of alleviating one’s own felt moral or physical contamination (e.g., physical cleansing, self-punishment, engaging in compensatory moral behavior) would eliminate this effect.
Although the present studies investigated how the punishment of a moral transgressor can ameliorate a salient individual-level threat to one’s personal moral identity, the historic purification rituals that inspired this research were notably collective in nature. Such rituals maintained the perceived moral purity of the greater social community by transferring the community’s sins onto a target that could be punished or expelled. Recent psychological research shows that people are strongly motivated to perceive the social groups with which they identify as moral (Leach, Ellemers, & Barreto, 2007). There is also a large body of research showing that people can feel guilt over illegitimate harm perpetrated by their group, even when they bear no personal responsibility for the harm done (see Branscombe & Doosje, 2004). This raises the possibility that the phenomenon investigated in the present studies in response to an individual-level morality threat may also operate in response to group-level morality threat when people self-categorize at the level of their group. If so, we would expect that reminding people of their group’s immoral actions should motivate people to see the group’s perceived moral contamination in target transgressors who can be punished to restore the group’s moral identity. Given that the threat is provoked at the level of the group, it is possible that a stigmatized subgroup, rather than an individual, may serve as the projection target under such conditions. Acts of “ethnic cleansing” may represent an extreme form of this behavior.

**Broader Theoretical and Practical Implications**

In addition to integrating and extending previous lines of research, the current studies also have broader implications about the motivation, function, and consequences of punishing moral transgressors. In particular, these studies help shed new light on people’s age-old fascination with punishment, ranging from public executions (Foucault, 1977) to Court TV. Much of the existing psychological literature suggests that the urge to see a transgressor
punished is grounded in the desire to see the world as a just place where people get what they deserve (Lerner, 1980). From this perspective, an unpunished transgressor poses a threat to one’s belief is a just world which is ameliorated by the transgressor’s punishment. This suggests that underlying the desire to see a transgressor punished is the desire to see justice done.

The current research offers a complementary perspective pointing to another motive underlying people’s desire to see a transgressor punished; the desire to expunge the taint of guilt left by their own unethical behaviors. This is consistent with Alexander and Staub’s (1956) analysis of criminology and penology claiming that individuals’ urge to see criminal offenders punished can reflect their own internal moral conflicts. Supporting this contention, recent research has shown that reminding people of their own group’s culpability for illegitimate harm-doing increases support for punishing another perceived harm-doer (Rothschild, Landau, Molina, Branscombe & Sullivan, in press). Importantly, the current studies go beyond this research by showing that the punishment of a moral transgressor alleviates people’s feelings of physical and moral contamination over their own immoral actions. This process may have particularly serious consequences in real-world contexts in which people make punitive judgments of others, such as in a court of law.

The current research also has important implications for people’s general willingness to engage in both prosocial and antisocial behavior. A litany of research has shown that threats to people’s moral identity can motivate compensatory efforts to engage in moral behaviors (e.g., Carlsmith & Gross, 1969; Darlington & Macker, 1966; Regan, Williams, & Sparling, 1972). The current research found that exposure to a punished transgressor bolstered people’s perceived personal moral identity in response to a moral value threat and reduced their willingness to engage in such compensatory behavior. This suggests that the punishment of a single moral
transgressor may serve to reduce the willingness of countless other people to engage in prosocial behavior.

Other research has shown that boosting people’s perceived moral self-concept can provide them with the “moral license” to engage in immoral behavior without experiencing aversive feelings of guilt (e.g., Merritt, Effron, & Monin, 2010). If exposure to a punished transgressor boosts people’s own perceived morality, this suggests that the punishment of a single moral transgressor may ultimately increase countless others’ willingness to engage in antisocial behavior. Thus, although institutions may seek to publicize the punishment of a moral transgressor to promote moral behavior and deter others from transgressing, in certain contexts, this type of exposure may inadvertently have the opposite effect. In a world saturated with sensationalized stories of crime and punishment, the present research highlights the potentially negative impact such information can have on people’s thoughts and actions.

**Conclusion**

The current studies provide strong evidence linking pre-modern rituals of atonement with a modern psychological phenomenon in which the punishment of a moral transgressor serves to morally cleanse the self. Of course, these studies also open up a slew of additional questions. Most notably, what are the boundary conditions for this phenomenon? Does the punished transgressor’s moral violation need to be more severe than one’s own? Is the cleansing effect equally potent when the punishment is more or less severe? Does the punishment need to be perceived to “fit the crime”? What would happen if people witnessed the moral transgressor administer self-punishment? Are there crimes for which this type of moral cleansing would be ineffective? Under what conditions might people choose this method of moral cleansing over other non-punitive guilt reduction strategies? Are there individual differences that may facilitate
or inhibit this cleansing effect? Can this effect be generalized to different cultural and religious contexts?

Importantly, the present studies do not reveal exactly why punishment is perceived to cleanse a moral transgressor. Is a punished transgressor seen as more moral because he or she is perceived to have learned his or her lesson, and is thus less likely to act immorally? Or is the punishment more broadly seen as a means of squelching deviance and bringing the transgressor back in line with social norms? Furthermore, reflecting on the larger social implications of this phenomenon, we might consider whether political leaders might strategically provide scapegoats to be punished as a means of giving their followers respite from their immoral actions in order to maintain social control? An investigation of this possibility could provide insight into horrific real world acts of “ethnic cleansing” or genocide. Future research targeting such questions will help to paint a more nuanced depiction of this phenomenon and its relevance in the real-world.
References


Footnotes

1. It is also worth noting that these findings provide what is, to our knowledge, the most direct evidence to date that people can feel soiled by their own moral violations. Zhong and Liljenquist (2006) indirectly assessing participant’s felt physical contamination by measuring the accessibility of cleaning-related words (Study 1), or willingness to pay more for cleaning products (Study 2). The current study directly assessed participants’ self-reported feelings of personal physical cleanliness.

2. Data for 11 participants was excluded from all analyses for failing to correctly identify that the new article described a hit-and-run car accident.
Table 1

*Ratings of target as “evil” and “dirty” as a function of target type (Supplemental Study)*

<table>
<thead>
<tr>
<th>Target Type</th>
<th>Evil</th>
<th>Dirty</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-Transgressor</td>
<td>2.78(_a) (1.39)</td>
<td>2.81(_a) (1.36)</td>
</tr>
<tr>
<td>Non-Punished Transgressor</td>
<td>3.95(_b) (1.33)</td>
<td>3.84(_b) (1.54)</td>
</tr>
<tr>
<td>Punished Transgressor</td>
<td>2.87(_a) (1.34)</td>
<td>2.98(_a) (1.22)</td>
</tr>
</tbody>
</table>

*Note.* Scale ranged from 1 to 6, with higher scores indicating descriptor was more characteristic of transgressor target. Means that do not share a subscript within the same column differ at \( p \leq .05 \).
Table 2

*Reported interest in joining blood drive as a function of personal immorality salience and target type (Study 2)*

<table>
<thead>
<tr>
<th>Personal Immorality Salience</th>
<th>Immorality Salient</th>
<th>Immorality Not Salient</th>
</tr>
</thead>
<tbody>
<tr>
<td>Target Type</td>
<td>No (48%)</td>
<td>Yes (52%)</td>
</tr>
<tr>
<td>Non-punished Transgressor</td>
<td>10</td>
<td>11 (52%)</td>
</tr>
<tr>
<td>Punished Transgressor</td>
<td>19 (86%)</td>
<td>3 (14%)</td>
</tr>
<tr>
<td></td>
<td>16 (76%)</td>
<td>5 (24%)</td>
</tr>
<tr>
<td></td>
<td>15 (65%)</td>
<td>8 (35%)</td>
</tr>
</tbody>
</table>
Figure 1. Ratings of personal physical cleanliness as a function of personal immorality salience and target type (Study 1).
Total Effect: $\beta = .28^{**}$
Direct Effect: $\beta = .19$ n.s.

Note: Total adjusted $R^2$ for the model = .10, $F(2, 84) = 6.00$, $p = .004$. All path coefficients represent standardized regression weights. The direct effect coefficient represents the effect of personal immorality salience on the dependent variable after controlling for the effect of the proposed mediator.

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

Figure 2. Indirect effect of personal immorality salience on perceived physical dirtiness of target transgressor through feelings of personal guilt (Study 2).
**Figure 3.** Ratings of personal morality relative to other students as a function of personal immorality salience and target type (Study 2).
Figure 4. Odds ratios of reported interest in joining a blood drive as a function of personal immorality salience and target type (Study 2).
Appendix A

Personal immorality salience primes used in Studies 1 and 2

Take a few moments to recall a time when you acted in an unethical way. This should be an action that you committed at least a year ago, but after your childhood.

Rather than think of a trivial action, think about a personally significant action. In the space below, write about this time. Describe how you acted unethically, what the consequences of your actions were, and how it made you feel about yourself. Your personal, honest responses are appreciated, and will be kept completely confidential. Continue writing until the computer presents you with the next task.

Writing prompt for personal immorality salient condition

Take a few moments to recall a time when you were very bored. This should be a time that happened at least a year ago, but after your childhood.

In the space below, write about this time. Describe what the situation was like, how it made you feel about yourself, and what actions you took. Your personal, honest responses are appreciated, and will be kept completely confidential. Continue writing until the computer presents you with the next task.

Writing prompt for personal immorality-not-salient condition
Appendix B

Student misconduct report, case description seen by all participants (Study 1)

<table>
<thead>
<tr>
<th>KU THE UNIVERSITY OF KANSAS</th>
</tr>
</thead>
<tbody>
<tr>
<td>MISCONDUCT REPORT SUMMARY</td>
</tr>
<tr>
<td>CATEGORY:</td>
</tr>
<tr>
<td>Student Disciplinary Hearing</td>
</tr>
<tr>
<td>REPORT STATUS:</td>
</tr>
<tr>
<td>Approved</td>
</tr>
</tbody>
</table>

**REPORT TITLE:**

Disciplinary Action Hearing Board Report Summary: Case #789621

**DATE:** [Redacted]

**PURPOSE:**

To describe and record findings of disciplinary hearing involving J. S. [Redacted] (referred to hereafter as J. S.).

**Case Description**

According to filed report #3185933, J. S. was accused of taking money from a donations box set up outside the Student Union. The donation box was set up by the organization Mothers Against Drunk Driving (MADD) for the purpose of collecting funds for the victims of drunk driving accidents.
Appendix C

Target type manipulation (Study 1)

**Final Judgment**

**Verdict: NOT GUILTY**

Upon hearing all evidence and testimony the board unanimously found the defendant to be not guilty of misconduct. Specifically, the investigation found no evidence that funds were stolen, and all donations were accounted for.

**Disciplinary Action**

**Disciplinary Action taken: NONE**

In accord with the Student Code of Conduct guidelines (Section 3, paragraph 5) no disciplinary action was taken.

*Non-transgressor condition*

**Final Judgment**

**Verdict: GUILTY**

Upon hearing all evidence and testimony the board unanimously found the defendant to be guilty of misconduct. Specifically, the investigation found that the defendant had stolen funds from a charitable organization on campus.

**Disciplinary Action**

**Action taken: NONE**

In accord with the Student Code of Conduct guidelines (Section 3, paragraph 5) misconduct was found, but no disciplinary action could be taken because the defendant was no longer a KU student.

*Non-punished transgresser condition*

**Final Judgment**

**Verdict: GUILTY**

Upon hearing all evidence and testimony, the board unanimously found the defendant to be guilty of misconduct. Specifically, the investigation found that the defendant had stolen funds from a charitable organization on campus.

**Disciplinary Action**

**Action taken: EXPULSION**

In accord with the Student Code of Conduct guidelines (Section 3, paragraph 5) misconduct was found and the student was expelled from the university.

*Punished transgresser condition*
Article introducing target transgressor (Study 2)

42-year-old woman left fighting for her life; hit-and-run driver mows her down

Posted: 11:23 p.m. EST, August 7, 2011

A careless motorist slammed into a 42-year-old woman while blowing through a red light in Orlando Wednesday night, witnesses and police said.

The driver ignored passerby’s pleas to stop after barreling into the woman - whose name was not released - and sped eastbound on Jackson St. after hitting the woman near S. Magnolia, cops and witnesses said.

“The driver was trying to beat the light,” said eyewitness Fernando Saez, 36, who was walking home when he saw a silver four-door sedan speeding along the avenue.

“The woman was trying to cross the street when the car hit her. The impact was so hard it sent the woman up in the air about 15-20 feet,” Saez said, still shaken from what he saw after the 6:20 p.m. accident.

Emergency responders found the victim lying on the ground at the scene, having difficulty breathing, authorities said. She was taken in serious condition with head and neck trauma to Lutheran Medical Center just eight blocks away where she was fighting for her life Wednesday night, police and fire officials said.

The driver of the silver Nissan - was identified earlier today as Thomas Carnahan, Jr., a 25-year-old student who lives in the area. The perpetrator’s whereabouts are currently unknown. Police have spent much of the week distributing photographs to local and state media asking the community to report any information that might lead to Carnahan’s arrest.
Appendix E

Target type manipulation (Study 2)

Driver still ‘at-large’ after hit-and-run accident that paralyzed woman

Posted: 8:55 a.m. EST, February 7, 2012

Six months ago Emma Nyman a 42-year-old pediatrician was hit by a car while walking across the intersection at Jackson and S. Magnolia.

Nyman survived the accident but sustained serious injuries leaving her paralyzed from the waist down.

The driver who fled the scene of the crime after hitting Nyman was identified by traffic cameras as 25-year-old Thomas Carnahan, Jr. a student at nearby Anthem Community College.

Local police issued a warrant for Carnahan’s arrest shortly after the incident, but have thus far failed to apprehend him.

If caught, Carnahan will likely face a series of criminal charges including one count of assault with a motor vehicle, one count of reckless endangerment, one count of leaving the scene of an accident without reporting as a felony, and resisting arrest.

If found guilty on all counts, Carnahan could serve up to 3 years in prison.

Article used in non-punished transgressor condition
Driver sentenced to 3 years of prison for hit-and-run that paralyzed woman

Posted: 8:55 a.m. EST, February 7, 2012

Six months ago Emma Nyman a 42-year-old pediatrician was hit by a car while walking across the intersection at Jackson and S. Magnolia Street.

Nyman survived the accident but sustained serious injuries leaving her paralyzed from the waist down.

The driver who fled the scene of the crime after hitting Nyman was identified by traffic cameras as 25-year-old Thomas Carnahan, Jr, a student at nearby Anthem Community College.

Local police issued a warrant for Carnahan’s arrest shortly after the accident and successfully apprehended him three days later.

Carnahan was convicted on numerous charges that included one count of assault with a motor vehicle, one count of reckless endangerment, one count of leaving the scene of an accident without reporting as a felony, and resisting arrest.

Carnahan is currently serving a 3-year prison term for his crimes.

Article used in punished transgressor condition
http://hdl.handle.net/1808/12199

*Share your story about how Open Access to this item benefits YOU at https://openaccess.ku.edu/you*