Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Psychology

Moral cleansing

Colin West[3,1] and Chen-Bo Zhong[2]

Moral cleansing describes behaviors aimed at restoring moral self-worth in response to past transgressions. People are motivated to maintain a moral self-image and to eliminate apparent gaps between their perceived self-image and their desired moral self. Moral cleansing behaviors fall into three over-arching categories. Restitution cleansing behaviors directly resolve past misdeeds. Behavioral cleansing involves counter-balancing across multiple dimensions of the moral self whereby threats in one subdomain are alleviated by bolstering a separate subdomain. Symbolic cleansing includes restitution behaviors that are only symbolically connected to the provoking moral threat, such as physical or ritual cleansing. The moral cleansing literature seeks to understand these seemingly erratic sequences of compensatory behaviors.

Addresses
[1] BEworks, 317 Adelaide Street West, Suite 400, Toronto, Ontario M5V 1P9, Canada
[2] Rotman School of Management, University of Toronto, 105 St. George Street, Toronto, Ontario M5S 3E6, Canada

Corresponding author: Zhong, Chen-Bo (Chenbo.zhong@rotman.utoronto.ca)
[3] Tel.: +1 416 400 1717.

Current Opinion in Psychology 2015, **6**:221–225

This review comes from a themed issue on **Morality and ethics**

Edited by **Francesca Gino** and **Shaul Shalvi**

For a complete overview see the Issue and the Editorial

Available online 3rd November 2015

http://dx.doi.org/10.1016/j.copsyc.2015.09.022

''Every saint has a past, and every sinner has a future.''
- Oscar Wilde

In a world of saints and sinners, predicting moral behaviors would be easy. The reality is that people are neither saints nor sinners, and ethical choices cannot be reduced to moral character alone. Moral behaviors are often influenced more by social context and by one's recent behavioral history than good versus evil. Much of ethical decision-making may be driven by a continuous struggle to maintain self-worth amidst our transgressions, trespasses, and temptations. Research on moral cleansing seeks to understand the myriad of behaviors that are observed in response to past transgressions. These seemingly erratic behaviors can only be understood within their temporal context, as direct and indirect responses to preceding ethical choices. This represents a stark departure from earlier traditions in moral psychology, which have primarily studied moral behaviors isolated in time, leaving the dynamics of sequential ethical decisions largely unexplored [1].

More recently, scholars have begun to examine situational influences on ethical decision-making and taken seriously the notion that ethical judgments and behaviors are shaped by their temporal location: their place within a series of choices and behaviors. Research on choice bracketing and temporal bracketing has examined the differences that arise between choices made in isolation versus those made as part of a series [2]. There is now an extensive literature exploring the cumulative and sequential effects of decision-making over time [3–5]. Thus, the phenomenon of moral cleansing is part of a burgeoning literature on moral self-regulation and compensatory ethics that examines sequential moral decisions and behaviors. Moral cleansing pays specific attention to how people respond to preceding misdeeds and moral threats. There are three over-arching categories of moral cleansing behaviors, each representing a different route toward restoring moral self-worth.

Categories of moral cleansing

Restitution cleansing

The most direct way to 'cleanse' oneself of a misdeed is simply to correct the wrong. Restitution cleansing is directly aimed at eliminating the provoking moral threat itself. Research by Tetlock, Kristel, Elson, Green, and Lerner [6] found that when violations of sacred values were envisioned, people were motivated to re-affirm these values. Contemplating a market for buying and selling human babies, or for organ trading, for example, led people to become more willing to participate in rallies against marketizing sacred values. Similarly, Dutton and Lake [7] found that giving white participants false feedback on a physiological measure indicating that they hold certain prejudices against racial minorities subsequently led them to give more money to a black versus white panhandler. Participants effectively engaged in reverse discrimination as a means of cleansing themselves of possible racist tendencies.

Behavioral cleansing

Many cleansing responses are far less direct than restitution cleansing. Restoring moral self-worth can involve counter-balancing across several disparate dimensions of moral character in ways that are usually opaque to outside

observers [8,9,10•]. In the moral cleansing literature these counter-balancing tendencies across subdomains can be described as 'behavioral cleansing.'

Early work by Carlsmith and Gross [11] found that people tend to be more socially cooperative after their moral values have been violated by participation in a mock version of the Milgram experiment. Similarly, people experience a stronger desire to engage in community service after receiving false feedback on reckless driving [12] and are more likely to donate to a homeless shelter after recalling past unsafe sexual activities [13]. These effects have since been characterized as a form of balancing moral 'debits' and 'credits' [14]. In this conceptual construct, moral assets are highly fungible. Debits accumulated from misdeeds in one moral domain can be offset by credits in a completely unrelated domain.

More recently, Gneezy et al. [15••] described this balancing process as 'conscience accounting,' and used this framework to explain specific preference reversals and time-inconsistencies in economic decision-making. Using an endogenous manipulation of moral self-image, they found that participants who chose to lie to increase their profits in a deception game subsequently donated more money to charity. In this case, participants accumulated a 'moral debt' by behaving dishonestly and paid for this offense by being more generous.

Symbolic cleansing

The third category of moral cleansing behaviors adds a new dimension of complexity to the narrative of dynamic moral behaviors as oscillating compensation and licensing, or moral debits and credits. Not only can restitution behaviors work across subdomains of the moral self, they can be purely symbolic and metaphorical.

Research in cognitive linguistics suggests that people frequently use conceptual metaphors to understand social interactions [16,17]. Conceptual metaphors may be particularly useful when facing abstract problems like those in moral judgments and decisions. The overlap between morality and bodily purity is a pervasive metaphor in everyday language–consider, phrases like 'getting your hands dirty' and 'having a dirty mouth.' There are several other parallels including similar facial expressions in response to moral transgressions and open wounds [18], and overlapping neural networks [19].

Building upon this metaphorical connection, several recent studies have found that physical cleansing can serve as a symbolic act of restitution that substitutes actual efforts to correct past wrongs. Zhong and Liljenquist [20•] asked participants to recall past transgressions and then gave half of them a chance to cleanse their hands using an anti-septic wipe. They found that physical cleansing literally reduced subsequent volunteer behavior. Further investigations found that this effect of physical washing is more pronounced among patients with Obsessive Compulsive Disorder (OCD) [21], and is grounded in the somatosensory cortex in the brain [22]. Similarly, religious rituals like confession may serve to symbolically 'cleanse' people of past sins, and potentially license future transgressions [23] (Ayal *et al.*, unpublished).

Moral cleansing theory posits that restitution, behavioral and symbolic cleansing are different manifestations of a common underlying mechanism. Thus, the explanatory and predictive strength of moral cleansing is grounded on two assumptions. First, a moral self-perception is desirable. And second, a moral self-perception is malleable in response to behaviors.

Two core assumptions of moral cleansing

People desire a moral self-image

Moral cleansing is only worthwhile if having a clean conscience is something desirable. People experience negative emotions when there is a decline or threat to their moral self-image caused by breaking established moral norms or recalling past transgressions [24]. On the other hand, people consistently report feeling a 'warm glow' and sense of satisfaction after behaving altruistically [25,26]. For instance, people report being happier after spending money on others compared to selfish spending [27].

It seems that a positive self-image is desirable and can often outweigh the lure of selfish gains. Mazar, Amir and Ariely's [28] study on the psychological underpinnings of dishonesty indicates that a desire to maintain a moral self-concept is strong enough to constrain selfishness and dishonesty. The desired moral self also serves as a reference point against which people measure their actual moral self at any given point in time [29]. It is distressing if there is a perceived comparative deficit. And while there are individual differences in the importance one assigns to maintaining a moral self-image [9], in general, people are motivated to eliminate any discrepancies.

Moral self-image changes in response to behaviors

How can people eliminate perceived gaps between their actual and ideal moral self? This question leads to the second core assumption of moral cleansing: that moral self-perception is linked to moral behavior. More specifically, the ethical implications of our behaviors, both real and symbolic, have the power to shift our moral self-perceptions. Immoral acts can depress our self-image and moral acts can boost our self-image.

Daryl Bem's work on self-perception theory [30] posits that people generally do not have a clear view of their own attitudes and internal states. It may be that we can only 'know' our internal states by observing our behaviors and then inferring that they must have been caused by these