



ELSEVIER

Contents lists available at ScienceDirect

Research in Organizational Behavior

journal homepage: www.elsevier.com/locate/riob



A dynamic and cyclical model of bounded ethicality

Dolly Chugh^{a,*}, Mary C. Kern^b

^a New York University, United States

^b Baruch College, United States

ARTICLE INFO

Article history:
Available online xxx

Keywords:
Ethics
Morality
Self
Automaticity
Bounded ethicality

ABSTRACT

We introduce a new model of bounded ethicality which helps explain three persistent puzzles of ethical behavior: when moral awareness is or is not present, when ethical behavior is more or less consistent with past behavior, and when blind spots obscure our ethical failures. The original conception of bounded ethicality (Chugh, Banaji, & Bazerman, 2005) described the systematic psychological constraints on ethical behavior and has contributed to our field's understanding of the phenomena of everyday, "ordinary" unethical behavior. In this more detailed model, we delineate these systematic processes and mechanisms and show how concepts of automaticity, self-view, and self-threat play critical roles in our ethical decision-making. The model describes distinct, asymmetric patterns of (un)ethical behavior and pinpoints the contingency which determines which pattern is more likely to unfold, including when we will trend to more or less automaticity and more or less ethical behavior. Our model integrates and synthesizes many of the key models and findings in recent behavioral ethics research into a single, overarching model of ethical decision-making, offering an anchor for new questions and a new realm of study.

© 2016 Published by Elsevier Ltd.

Contents

| | |
|--|-----|
| 1. What is bounded ethicality? | 000 |
| 2. The field of behavioral ethics | 000 |
| 2.1. Two themes | 000 |
| 2.2. Three puzzles | 000 |
| 2.2.1. Puzzle #1: When do we see the ethical implications of decisions and when don't we? | 000 |
| 2.2.2. Puzzle #2: Does good behavior generate more good behavior, or license bad behavior? | 000 |
| 2.2.3. Puzzle #3: Why are we sometimes blind to our own transgressions? | 000 |
| 3. The bounded ethicality model | 000 |
| 3.1. Summary of how our model works | 000 |
| 3.2. Key psychological concepts | 000 |
| 3.2.1. Self-view | 000 |
| 3.2.2. Self-threat | 000 |
| 3.2.3. Self-enhancement and self-protection | 000 |
| 3.3. How the model works | 000 |

* Corresponding author.
E-mail addresses: dchugh@stern.nyu.edu (D. Chugh), mary.kern@baruch.cuny.edu (M.C. Kern).

| | | |
|--------|---|-----|
| 3.4. | How the model reconciles puzzles in the literature. | 000 |
| 3.4.1. | Puzzle #1: When do we see the ethical implications of decisions and when don't we?. | 000 |
| 3.4.2. | Puzzle #2: Does good behavior generate more good behavior, or license bad behavior? | 000 |
| 3.4.3. | Puzzle #3: Why are we sometimes blind to our own transgressions? | 000 |
| 4. | An agenda for future research. | 000 |
| 4.1. | Ethics research | 000 |
| 4.2. | Self research | 000 |
| 5. | Conclusion | 000 |
| | References | 000 |

1. What is bounded ethicality?

The original model of bounded ethicality challenged the notion that people can be fully ethical all the time, proposing instead that we are all prone to ethical failure (Chugh, Banaji, & Bazerman, 2005). Further, these ethical failures are the outcome of systematic and ordinary psychological processes (Chugh et al., 2005), and these processes are neither rare nor unpredictable (Greenwald & Banaji, 1995). Scholars using this model of bounded ethicality (e.g. Bazerman & Tenbrunsel, 2012) have focused primarily on the ways in which bounded ethicality helps explain phenomena in which a gap exists between our intended and actual ethical behavior. For example, within organizations, bounded ethicality explains a wide range of phenomena in which we are not as ethical as we think we are, plan to be, or want to be (Tenbrunsel & Smith-Crowe, 2008). We overclaim credit (Caruso, Epley, & Bazerman, 2006), discount the future (Wade-Benzoni, Tenbrunsel, & Bazerman, 1996), and are swayed by conflicts of interest (Moore, Cain, Lowenstein, & Bazerman, 2005).

While the bounded ethicality model has been heavily cited by both researchers and practitioners, its original conception contributes more to our understanding of the phenomenon represented by bounded ethicality, and less to our understanding of the psychology underlying bounded ethicality. In this paper, we advance and reform the “1.0” model (Chugh et al., 2005) with a far more comprehensive “2.0” model. The primary contribution of our new model lies in the expansion and specification of the underlying psychological processes of bounded ethicality. We focus less on what bounded ethicality explains, and more on how to explain bounded ethicality.

Boundedness is a powerful concept that has been summoned by many scholars for a range of purposes. The scholarly utility of the boundedness concept began, of course, with Simon (1957). He enlisted this notion of boundedness to describe the limited information available to decision-makers, the limited cognitive capacity of decision-makers’ minds, and the limited time available for decision-makers, all of which bound the rationality of the decision itself. The descriptive power of boundedness lay not only in its accuracy in capturing limitations on rationality, but also in the rich literature which soon emerged on heuristics and biases, delineating the systematic nature of those limitations (Tversky & Kahneman, 1974). Boundedness, thus, came to carry meaning of not just limitations, but systematic limitations (i.e. limitations that are predictable and non-random).

As the behavioral ethics field has gained research momentum, the notion of boundedness has become important and useful, both for our work and the work of others.¹ The original work on bounded ethicality (Chugh et al., 2005) leveraged this important idea in defining bounded ethicality as the systematic and ordinary psychological processes that constrain the ethicality of decision-making. Thus, bounded ethicality both named the distinction between our unbounded ethicality and the reality of our behavior, and specified that our behavior is driven by a systematic set of processes.

In this paper, we revise this definition of bounded ethicality to be: *the systematic and ordinary psychological processes of enhancing and protecting our ethical self-view, which automatically, dynamically, and cyclically influence the ethicality of decision-making.* This revised definition retains the emphasis on “systematic and ordinary”. But, importantly, it does not assume that bounded ethicality always constrains ethicality, thus allowing for bounded ethicality to explain both ethical and unethical behavior, which is a significant shift from the original conception of bounded ethicality. The new definition also specifically notes the roles of the self and automaticity and highlights the automatic, dynamic, and cyclical features of the process, thus delineating underlying processes which were not previously captured.

A key contribution of our model is its integration of existing literatures and theories in the behavioral ethics literature into a single model of ethical decision-making. The bounded ethicality model aligns well with a wide and seemingly disparate range of important perspectives of the past ten years. Many of these models of ethical decision-making lean heavily to either the motivational or the cognitive perspective. We believe that bounded ethicality is the first to balance and integrate both perspectives, providing the behavioral ethics field with a starting point for future research on how these perspectives work together.

This paper is organized as follows. We provide the reader with developments in the recent explosion of behavioral ethics research, laying out two key themes and

¹ The notion of boundedness has been orbiting ethics research in multiple forms in recent years, both predating and postdating the introduction of bounded ethicality. Several terms that sound similar to bounded ethicality exist, but mean different things. In addition, several terms are distinctly different from bounded ethicality but carry similar meaning. For the interested reader, a comprehensive list of these terms is provided in Table 1.

Table 1
Concepts related to ethics and boundedness.

| Concept | Citation | Distinction from bounded ethicality |
|--|---|---|
| Bounded moral rationality | Donaldson and Dunfee (1994) | Narrowly focused on the moral challenges of the economic context, not the broader psychological context. |
| Bounded rationality of implicit social beliefs | Banaji and Bhaskar (2000) | <p>Focused on one specific domain of mental processing (stereotyping), not ethical decision-making more broadly.</p> <p>More similarly, Banaji and Bhaskar (2000) argue that the bounded rationality of human social cognition and implicit social beliefs best characterizes the processes of stereotyping and prejudice. They challenge notions of stereotyping which treat the process as rational or as limited to a few individuals, and rather, present stereotyping both as boundedly rational and ordinary. The bounded rationality of implicit social beliefs is thus a subset of bounded ethicality, capturing the same elements but in one specific domain of mental processing (stereotyping) and with an orientation leaning towards the cognitive perspective, while bounded ethicality captures both a cognitive and motivational perspective.</p> |
| Bounded personal ethics | Murnighan, Canteloni, and Elyashiv (2001) | <p>Focuses on ethical awareness as a non-automatic process and self-interest as an automatic process, not on the impact of self-view.</p> <p>Describes how individuals are often unaware of the ethical implications of their actions. In this model, awareness of ethical implications does not emerge automatically, but self-interest does. Thus, in the absence of ethical awareness, the automatic process of protecting one's self-interest dominates (whereas, in bounded ethicality, the automatic process of enhancing and protecting one's self-view dominates).</p> |
| Bounded ethical awareness | Gino and Bazerman (2009) | Refers to the tendency to overlook the unethical behavior of others when it develops gradually over time, rather than abruptly at once. Helpful in understanding our moral judgments of others, but does not relate to one's own ethical decision-making directly (though one's judgments have implications for one's willingness to address or report the unethical behavior of others). |
| Moral heuristics | Sunstein (2005) | <p>Focuses on perceptions of others' behavior, not one's own behavior.</p> <p>Sunstein (2005) provides examples of beliefs that people hold, such as "punishment should be commensurate with outrage", and demonstrates that this belief functions as a moral heuristic (mental shortcut) that works well in most situations but sometimes generates "absurd" moral judgments with serious implications for law and policy. Like bounded ethicality, moral heuristics refer to the automaticity aspect of mental processing, but unlike bounded ethicality, moral heuristics are most relevant to the judgments of others rather than one's own behavior.</p> |
| Social decision heuristics | Allison and Messick (1990) | <p>Focuses on situations of shared resources.</p> <p>Social decision heuristics are shortcuts that are applied to a subset of situations in which bounded ethicality is relevant: situations where people share a common resource and must determine their individual levels of resource consumption (Allison & Messick, 1990). These various heuristics are not synonymous with bounded ethicality, but align well with bounded ethicality by capturing a systematic and ordinary automatic psychological process which affects ethical decision-making or ethical judgment.</p> |
| Responsibility gap | Margolis (2009) | <p>Describes "the discrepancy between the ethical responsibilities managers must take upon themselves ... and the realities of human functioning that limit our human capacity to meet those responsibilities".</p> <p>While the concept does not explicitly reference boundedness, it reflects its essence in describing the disconnect between the complex ethical responsibilities facing managers (often because the responsibilities are not being assumed by other individuals or institutions) and the limitations of human functioning (Margolis, 2009). Like bounded ethicality, the responsibility gap incorporates the reality that humans may face situations with intensely challenging ethical implications, and that one's intentions and one's actual behavior may not align. As a result, "even good intentions and good systems may be overmatched at times by situational pressures, social dynamics, and psychological tendencies, leaving responsibilities unfulfilled" (Margolis, 2009: 42). The responsibility gap thus represents a specific manifestation of bounded ethicality, seen in extreme situations of overwhelming ethical challenge.</p> |

three persistent puzzles. We then import several important concepts from the psychological literature, so that we may use them to build out our model. With this background in place, we delineate our model, focusing on the processes of self-enhancement and self-protection, and how they hinge on the process of self-threat assessment, resulting in asymmetric patterns of (un)ethical behavior. We weave in real-world examples to make our model concrete for the reader. Finally, we discuss the implications of our bounded ethicality model for both future research and practice.

2. The field of behavioral ethics

Behavioral ethics research is thriving, revealing pattern after pattern of surprising and counter-intuitive evidence that we are not perfectly or consistently ethical all of the time. In this section, we will highlight two conceptual themes in the literature and describe three puzzles which persist, or even emerge, as this research grows. These puzzles persist despite the robust set of empirical findings and the impressive array of other integrative models in the field (Haidt, 2008; Jones, 1991; Monin & Jordan, 2009; Reynolds, 2006a; Sonenshein, 2007; Tenbrunsel, Diekmann, Wade-Benzoni, & Bazerman, 2010; Tenbrunsel & Smith-Crowe, 2008; Treviño, 1986; Zhong, Ku, Lount, & Murnighan, 2010; Zhong, Liljenquist, & Cain, 2009).

We posit that our model of bounded ethicality addresses these persistent puzzles in a way that is novel and distinctive from other existing models. We begin by describing the themes and puzzles themselves and will later show how our model helps resolves these puzzles.

2.1. Two themes

Two themes recur across researchers and methods: the role of automaticity and the role of the self. In the original conception of bounded ethicality (Chugh et al., 2005), much of our unethical behavior takes place outside of our awareness. Similarly, recent advances in behavioral ethics explore the role of automaticity in ethical decision-making, challenging rationalist models (Haidt, 2002; Reynolds, 2006a; Reynolds, Leavitt, & DeCelles, 2010; Sonenshein, 2007; Welsh & Ordóñez, 2014). Other work focuses on the impact of conditions fostering automaticity on ethical decision-making, such as time pressure, cognitive load, and depletion (e.g. Gino, Schweitzer, Mead, & Ariely, 2011; Kern & Chugh, 2009; Mead, Baumeister, Gino, Schweitzer, & Ariely, 2009).

The second theme centers on the self. Whether referred to as one's "self-image" (e.g. Zhong et al., 2010), "self-view" (e.g. Miller & Effron, 2010; Monin & Jordan, 2009), "self-concept" (e.g. Mazar, Amir, & Ariely, 2008), or "moral self" (e.g. Jordan, Mullen, & Murnighan, 2011; Zhong et al., 2009), many researchers have measured or manipulated the degree to which an individual feels secure in their view of themselves. Ethical behavior seems to pivot around one's self-view, generating a fascinating set of what appears to be contradictory findings.

2.2. Three puzzles

Despite the emergence of these two clear themes, we still have many questions about what shapes ethical behavior. Three puzzles, in particular, stand out.

2.2.1. Puzzle #1: When do we see the ethical implications of decisions and when don't we?

Most of the time, individuals do not use an ethical frame for decision-making (Tenbrunsel & Messick, 1999). And most decision-making is accompanied by a general low level of moral awareness (Tenbrunsel & Smith-Crowe, 2008), such that the individual fails to interpret a situation as containing ethical issues or as relevant to moral principles (Butterfield, Treviño, & Weaver, 2000; Ferrell & Gresham, 1985; Rest, 1986; Reynolds, 2006b). Without moral awareness, amoral decision-making can result in either ethical or unethical choices. But sometimes, we do operate with moral awareness and when we do, ethical decisions are more likely (Tenbrunsel & Smith-Crowe, 2008).

So, our model tackles the puzzle of how and when moral awareness is activated: when do we see the ethical implications of decisions and when don't we?

2.2.2. Puzzle #2: Does good behavior generate more good behavior, or license bad behavior?

The behavioral ethics literature has seen a surge of models reflecting dynamic processes.² What is clear in these models is that prior deeds (actual or even imagined) affect current deeds. But this work is also puzzling, leaving us to wonder whether being good begets future good behavior or begets future bad behavior. For example, sometimes, a prior good deed leads to another good deed as one builds a moral muscle (Baumeister, Vohs, & Tice, 2007), or a prior good deed leads to a bad deed (Effron & Monin, 2010; Miller & Effron, 2010; Monin & Jordan, 2009; Monin & Miller, 2001). In both cases, a dynamic process unfolds in which past behaviors shape one's self-view, which in turn shapes one's current and future behaviors – but it is unclear what determines the direction of this effect.

Relatedly, when do our unethical choices snowball to more ethically charged and more ethically significant transgressions? In other words, when do our unethical behaviors increase in moral intensity (Jones, 1991)? Our ordinary and small unethical behaviors can lead to far less ordinary and far more significant unethical behaviors over time. This cycle is akin to the slippery slope phenomenon (Gino & Bazerman, 2009; Gino, Moore, & Bazerman, 2009) in which people are less likely to respond to others' unethical behavior when it degenerates gradually rather than in a single abrupt shift. What makes this cycle most pernicious is that subsequent, consecutive unethical acts are unlikely to be of smaller proportion than the previous

² For example: the "fudge factor" theory (Mazar et al., 2008); moral self-regulation (Zhong et al., 2009); dynamic moral self (Monin & Jordan, 2009); psychological license (Miller & Effron, 2010); compensatory ethics (Zhong et al., 2010); moral equilibrium (Jordan et al., 2011); temporal selves (Tenbrunsel et al., 2010); and conscience accounting (Gneezy et al., 2014).

acts. Sometimes, subsequent acts are likely to be of greater moral intensity than previous acts, and we slide down the slope towards bigger and bigger transgressions.

So, our model of bounded ethicality tackles this puzzle of consistency versus inconsistency in the ethicality and moral intensity of ethical behavior: when does behavior lead to more of the same versus a shift?

2.2.3. Puzzle #3: Why are we sometimes blind to our own transgressions?

When one is able to engage in unethical behavior without feeling unethical, one might actually be blind to this disconnect. Bazerman and Tenbrunsel (2012) refer to this tendency to not see one's own unethicality as a "blind spot". For example, after committing an unethical act, individuals are more likely to forget moral rules prohibiting that act (Shu & Gino, 2012). Similarly, after committing an unethical act that harms someone else, individuals are more likely to morally disengage by construing the victim as bearing some responsibility for the harm (Bandura, 1986; Detert, Treviño, & Sweitzer, 2008). The desire to be good (versus not bad) combined with common psychological biases such as the confirmation bias (Einhorn & Hogarth, 1978) and the egocentric bias (Babcock & Loewenstein, 1997) leads to a self-view that is more favorable than it is accurate. These are instances of blind spots.

So, our model of bounded ethicality tackles this third puzzle related to why we sometimes see our own ethical failures, and at other times we are not able to see them: that is, *why are we sometimes blind to our own transgressions?*

3. The bounded ethicality model

With these puzzles in mind, we present our model. It is worth noting that much of the scholarly narrative on ethical failures, particularly prior to the most recent decade, is dominated by the role of self-interest and its rational underpinnings (Miller, 1999), though this premise is typically unspoken (Epley & Caruso, 2004; Moore & Loewenstein, 2004; Tenbrunsel & Smith-Crowe, 2008; Welsh & Ordóñez, 2014). We can infer this premise because

this type of ethics research often describes the importance of the individual or organization tamping down self-interest via levers as wide-ranging as moral development (Kohlberg, 1969), organizational climate (Victor & Cullen, 1988), norms (Pitesa & Thau, 2013; Treviño, 1986), and regulatory resources (Mead et al., 2009).

Bounded ethicality defies this self-interest-driven narrative of ethical decision-making, in which unethical behavior reflects the unconstrained pursuit of self-interest, while ethical behavior emerges from the conscious constraint of that self-interest. In bounded ethicality, self-view is a more forceful and more automatic influence than self-interest on ethical decision-making. In other words, bounded ethicality argues that self-view – that is, one's interestedness in one's self rather than one's self-interest – is the key to understanding how ethical behavior emerges in individuals and organizations.

This shift is both critical and timely, as the theoretical dominance of self-interest has waned in light of research on prosocial behavior (e.g. Grant, 2007), altruism (e.g. Fehr & Fischbacher, 2003), self-interest as a norm (e.g. Miller, 1999), and bounded self-interest (e.g. Jolls, Sunstein, & Thaler, 1998), all of which reveal the many occasions in which self-interest fails to adequately explain human behavior. We argue that self-view dominates self-interest. Self-interest motivates some behavior some of the time, but in our model, the impact of self-interest on ethical decision-making is likely bounded by the self-view.

In this section, we first lay the planks for the model with an overview of how it works. We develop a workplace example to facilitate building in the details of the model. We next describe several central psychological concepts – including self-view – which are central to our model. Then, we delve into the details of the model and use it to make sense of the three puzzles in the behavioral ethics literature.

3.1. Summary of how our model works

Our model is portrayed in Fig. 1, and we encourage the reader to take visual note of a few central features. First,

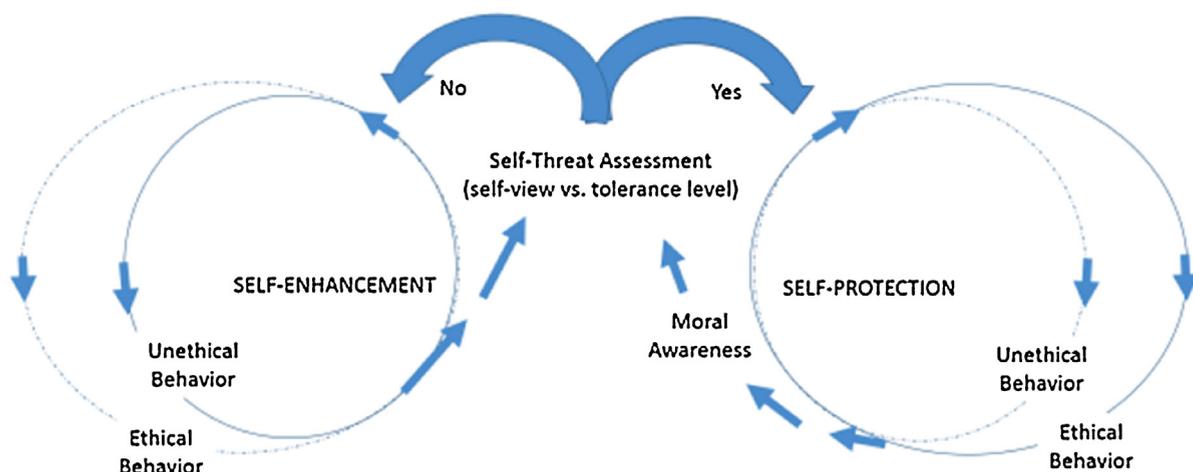


Fig. 1. Bounded ethicality model.

the model is asymmetric, with a different set of processes unfolding on the right side of the figure versus the left side of the figure. The details of this asymmetry will prove important as we unpack the model. Second, the model is dynamic because subsequent events pivot on what happens in the middle of the figure. This contingency determines whether the individual's behavior will flow to the right side versus the left side of the model, suggesting that ethical behavior does not unfold identically from occasion to occasion, even within the same individual. Third, the model is cyclical, suggesting that there is no obvious starting or ending point in how it unfolds in reality (or how it should be explained in writing).

As an overview, our model represents ethical behavior as hinging on an assessment in the middle of the figure – *self-threat assessment* – which determines whether either *self-enhancement* (the process depicted on the left of the model) or *self-protection* (the process depicted on the right of the model) processes will follow. Self-threat assessment is the process of determining whether an individual's self-view as an ethical person is threatened. Based on this assessment, either the process of self-enhancement (lower self-threat) or self-protection (higher self-threat) will follow. Self-enhancement and self-protection have specialized functions: self-enhancement processes increase the positivity of the self-view, while self-protection processes decrease the negativity of the self-view (Alicke & Sedikides, 2009; Sedikides, 2012). We provide extensive detail on these concepts next.

3.2. Key psychological concepts

For our model, we import several concepts from outside the ethics literature, whose concepts offer significant explanatory power in the ethics context: self-view, self-threat, self-enhancement, and self-protection.

3.2.1. Self-view

A fundamental tenet of the self literature is that we care deeply about our self-concept and self-esteem – known collectively as our self-view (Sedikides, 2012) – so much that people show a preference for heightening one's self-esteem over “eating a favorite food, engaging in a favorite sexual activity, drinking alcohol, receiving a paycheck, or seeing a best friend” (Bushman, Moeller, & Crocker, 2011). Thus, we are motivated to both protect and enhance our self-view (Alicke & Sedikides, 2009), even when doing so runs counter to our self-interest (in a utilitarian or economic sense).

Self-view in the moral domain is unique in two ways: the pervasiveness of a desire for a positive ethical self-view and the primacy of the domain in one's view of self. First, we are motivated to see ourselves as ethical and we typically rate ourselves as above average in ethicality, honesty, trustworthiness, and fairness (Baumhart, 1968; Epley & Caruso, 2004; Messick & Bazerman, 1996; Tenbrunsel, 1998). This desire is captured in the moral identity literature. Moral identity is the extent to which an individual perceives morality as central to his or her self-view (Aquino & Reed, 2002; Blasi, 2004). While the centrality of moral identity varies (Aquino & Reed, 2002;

Blasi, 2004; Pitesa & Thau, 2013), the vast majority of people report a desire for a positive ethical self-view (Higgins, 1987; Mazar et al., 2008; Nisan, 1991). The need for a positive self-view, and verification of this self-view from others, is not necessarily true for all domains. When individuals hold negative self-views in particular domains, they seek validation of these views from others, or “self-verification” (Swann, 1983). Self-verification theory recognizes that even validation of negative self-views may also be valued, under some conditions. But, our model of bounded ethicality is focused on the moral domain, specifically, and we argue that this domain is uniquely unlikely to be one with a negative self-view, given the pervasive and essential nature of positive ethical self-views. In cases where the individual holds a positive self-view, which is likely to be the case with ethical self-view, then self-verification and self-enhancement align such that it is unlikely that an individual will be seeking validation from others for his/her view of him/herself as unethical.

The second way in which the moral domain is unique lies in the primacy of the domain in one's view of self. Not only is the desire for a positive ethical self-view pervasive, it is also “essential” (Strohinger & Nichols, 2014). That is, moral traits are privileged in how we view the self, such that “moral traits are considered more important to personal identity than any other part of the mind” (Strohinger & Nichols, 2014: 168).

3.2.2. Self-threat

A challenge to an individual's self-view is known as self-threat, “when favorable views about oneself are questioned, contradicted, impugned, mocked, challenged, or otherwise put in jeopardy” (Baumeister, Smart, & Boden, 1996: 8), and can arise from “perceived or real doubt, diminishment, or devaluation” to the self-view (Leary, Terry, Allen, & Tate, 2009). Self-threats arise when one's self-view falls below one's tolerance level (Alicke & Sedikides, 2009). Although we rarely achieve our ideal level of functioning or aspirational level of being fully ethical, nor perhaps do we see ourselves objectively as others see us, self-threats emerge when one's negative ethical self-view is below the tolerance level, such that there is an unacceptable degree or frequency of unethical behavior (Alicke & Sedikides, 2009).

Self-threats leave individuals in a psychologically uncomfortable state, and the return to a positive ethical self-view can be achieved via either primary or secondary control mechanisms (Rothbaum, Weisz, & Snyder, 1982). Primary control mechanisms are behaviors in which the individual acts to change the objective state of affairs, such that a real difference is enacted which furthers the person's interests (Rothbaum et al., 1982); in our context, this would mean behaving more ethically. But primary control mechanisms are not the only way to have positive ethical self-view; one can also turn to secondary control mechanisms.

Secondary control mechanisms are substitutes for primary control behaviors; these are psychological processes that replace the behavior, but allow the individual to enjoy the same psychological benefits of a primary control behavior (Rothbaum et al., 1982). So, one can maintain a

positive ethical self-view by substituting a shift in how one thinks about one's behavior in place of actually behaving ethically. For example, a primary control strategy is to provide customers with a safe product; a secondary control strategy is to provide customers with safety warnings, justifying that any harm that comes to the user is due to his or her own negligence. In either case, primary or secondary control, the individual emerges with ethical self-threat effectively warded off. Loosely said, primary control behaviors in the ethics domain can be characterized as “be good” behaviors, while secondary control mechanisms can be characterized as “feel good, without being good”. In short, when an individual experiences a self-threat in the moral domain, the individual can turn to either primary or secondary control strategies to reduce the self-threat, and they can engage in these strategies in either the self-protection or self-enhancement processes.

3.2.3. *Self-enhancement and self-protection*

Self-enhancement and self-protection (Sedikides, 2009) are both oriented toward the goal to maintain a positive ethical self-view. The difference between self-enhancement and self-protection is in how this goal is accomplished. A critical difference between the two sets of the processes lies in the degree of automaticity.

That is, in order for individuals to not self-enhance, heightened dorsomedial frontal cortex activity is needed, suggesting that self-control is required to stop the otherwise-automatic self-enhancement processes (Krusemark, Campbell, & Clementz, 2008). Automatic self-enhancement has been seen throughout a wide range of literatures in psychology, such as slower response times for endorsing negative traits than positive traits (Paulhus & Levitt, 1987), preferences for letters that appear in one's name versus letters that do not (Nuttin, 1985), tendencies to apply stereotypes to others after receiving negative feedback about oneself (Gilbert & Hixon, 1991), and heightened self-enhancement after engaging in depleting activities (Vohs, Baumeister, & Ciarocco, 2005). This small sub-set of examples illustrates the ubiquity of self-enhancement and the automaticity with which it occurs.

In contrast, self-protection is intended to ensure that the individual's self-view is “not-bad” and occurs in response to episodic meaningful self-threats which evoke the potential for a negative self-image – “the greater the threat to the self, the more self-serving the attributions will be” (Campbell & Sedikides, 1999: 26). When a meaningful self-threat occurs, it has the potential to interrupt the automatic functioning more characteristic of self-enhancement until the positive self-view is restored. This is a more deliberative and conscious process.

The desire to be “good” or a desire to be “not bad” is akin to the ways in which regulatory focus theory describes a tendency to move towards a goal (promotion focus) versus a tendency to move away from the reverse of the goal (prevention focus) (Higgins, 1997, 1998; Zhang, Cornwell, & Higgins, 2014). Regulatory focus theory describes what is motivating the individual. Self-enhancement and self-protection describes the psychological process enacted in this pursuit.

3.3. *How the model works*

We now illustrate how these constructs interact through the lens of bounded ethicality. We define the “perceived level” of our positive ethical self-view as the point at which one believes one functions (this is what we refer to as the self-view) and can be at any point on the continuum, ranging from the aspiration level to below the tolerance level (Alicke & Sedikides, 2009). The aspiration level in the ethics domain can be conceptualized as full, or unbounded, ethicality; for example, this would be someone whose ethical behavior is completely unswayed by conflicts of interest, in-group favoritism, implicit bias, or social pressures. While this level of ethical perfection is moral ground few (none) can claim, one's perceived level (aka self-view) may nonetheless approach the aspiration level. Or, one's self-view may, at times, be closer to the tolerance level, such as when one capitulates to a social norm that allows for taking home office supplies for personal use (e.g. school supplies in August).

As shown in Fig. 1, self-threat assessment leads to a process of either self-enhancement or self-protection, contingent upon whether a self-threat exists. The self-threat is the difference between one's perceived level (one's self-view) and the tolerance level. When one's perceived level is below the tolerance level, a self-threat exists and one is motivated to restore one's perceived level to at least the tolerance level; this state of affairs triggers self-protection processes. Critically important, this less automatic set of processes also activates moral awareness, which makes self-protection less prone to the types of defensiveness in information processing that one would expect in self-enhancement.

In contrast, when one's perceived level is above the tolerance level, but short of the aspiration level, a self-threat does not exist and thus, one continues to be motivated to raise one's perceived level towards the aspiration level; this state of affairs triggers self-enhancement processes. Thus, absent ethical self-threat, self-enhancement processes are continually operating, while ethical self-threat triggers an episodic shift from self-enhancement to self-protective processes. Even when facing minimal self-threat, one's self-view is sustained on an ongoing basis through the continual operation of self-enhancement processes (Alicke & Sedikides, 2009), which are able to easily dispense of threats through either primary or secondary control strategies.

The processes are occurring over and over again, with no beginning or end point. In Fig. 1, the cyclical nature is revealed in the arrows that continually and repetitively bring the process around. The cycling of our model through self-threat assessment creates an ongoing pattern of further cycling through self-enhancement and/or self-protection.

However, while we present a dual cycle model, we are intentional that this is not a dual process model in which one process is a more routinized version of the other process. In a dual process model, two similar processes occur in parallel, perhaps at different speeds. A dual cycle model is one in which the two processes are distinctly different from each other, both in substance and perhaps

speed. As Reynolds (2006a: 741) describes dual-cycle models, “the two cycles process ethical situations in completely unique fashions, and therefore each has its own unique properties and implications for ethical decision making”. These two processes of self-protection and self-enhancement are asymmetric in their functioning and consequences, as we summarize in Table 2, and the asymmetry between these cycles is core to our model.

As a means of explaining our model, we will use a workplace example featuring a sales manager. She starts her day without meaningful ethical self-threat, and thus is able to sustain her usual positive ethical self-view. She receives two emails from customers this morning. One large buyer wants to give her a heads up that a noteworthy order might be cancelled due to budgetary issues; the decision will be made next week but he wanted to let her know for her planning purposes. This bad news is followed by potentially good news. A prospect, whom she has been wooing for months, expresses optimism that the needed approvals will come through so that they can finally place their first order soon.

She also receives an email from one of her direct reports containing the sales forecast that will be submitted today, confirming that she does not have any more revisions. This forecast will be seen by investors and by the sales manager’s boss who will use it as one data point in computing her performance bonus. Her direct report also asks if she would like to present the forecast during today’s sales team meeting, or if he should be prepared to do it. Finally, the last email in her inbox is from some former-coworkers-turned-friends; they are confirming plans to meet up for dinner later and catch up on life.

There are several different ways in which the sales manager might behave in these circumstances and we will use our model to explain which counterfactual actually occurs. Specifically, we will consider whether she downgrades the sales forecast to reflect the pending cancellation (more ethical), upgrades the sales forecast to reflect the possible sale (less ethical, if the cancellation is not also reflected), and expenses the dinner with former colleagues as a work-related expense (less ethical).

A critical determinant, of course, is the degree to which she experiences self-threat at various points in her day. Self-threat assessment is partially contingent upon both situational and dispositional variables. We can delineate a

number of moments in which her self-threat assessment might lead to a shift from self-enhancement to self-protection. To begin with, dispositional variables, such as a more central moral identity (Aquino & Reed, 2002), increase the likelihood that she will experience self-threat throughout her day, because her overall aspiration level will be relatively high and her tolerance level will be relatively low. So, to the extent that the sales manager values morality as central to her self-view, she may generally be more aware of ethical temptations.

And, individuals in situations that activate one’s moral identity are more likely to experience self-threat because their desire to be ethical is more salient (Welsh & Ordóñez, 2014). So, if the organizational culture (Treviño, 1986) the sales manager works in places great emphasis on integrity in the sales forecasts, this situational variable increases the probability of self-threat. We could also imagine if the sales forecast requires her signature and if it must be accompanied by a form which she signs before the final numbers can be filled in, not after, this situational variable of salience of honesty norms (Shu, Gino, & Bazerman, 2011) will influence her self-threat assessment.

We can also imagine that the sales manager’s self-view could be influenced by the language used by her and those around her. The influence of language is revealed through a simple manipulation of whether cheating behavior is described using nouns or verbs (Bryan, Adams, & Monin, 2013). Individuals cheated more when instructions referred to cheating using a verb (“cheating”) versus a predicate noun (“being a cheater”). This finding is consistent with work in the self-literature in which the degree to which a task is diagnostic of one’s self-concept moderates the degree of self-threat the individual experiences (Campbell & Sedikides, 1999), with nouns carrying more diagnostic cues than verbs. In our model, the explanation for this finding is that self-threat rises in the noun-based description. If our sales manager is in an environment where people who fudge the forecast and expense non-work-related meals are referred to as “cheaters”, she is more likely to experience self-threat, than if such behaviors are referred to as “cheating”.

In summary, the process of self-threat assessment is one in which the individual is continually re-assessing his or her behavior in order to determine whether a self-threat exists, with both situational and dispositional factors influencing this process for our sales manager. An ethical self-view below the tolerance level generates a self-threat, whereas an ethical self-view above the tolerance level is not a self-threat. The next step in our dynamic and cyclical model is either self-enhancement or self-protection, contingent on whether she faced a self-threat.

In Fig. 1, the dynamic nature of bounded ethicality is illustrated because what happens next, in terms of one’s ethical or unethical behavior, hinges largely on what has just happened in self-threat assessment. That is, it is dynamic because (1) one’s behavior is a function of one’s past and considered behaviors and (2) the mechanism for this relationship between the past and present lies in the person’s self-view. In our example, when our sales manager faces the forecast decision, our dynamic model predicts that her decision-making process will be partially

Table 2
 Asymmetry between self-enhancement and self-protection cycles.

| Self-enhancement cycle | Self-protection cycle |
|--|--|
| Minimal ethical self-threat Oriented towards the goal of a positive ethical self-view | Meaningful ethical self-threat Oriented towards the goal of a not-negative ethical self-view |
| More automatic processing Continual More blind spots | Less automatic processing Episodic, prompted by self-threat Fewer blind spots |
| Cycles towards unethical behavior over time | Cycles towards ethical behavior over time |
| Behavior is relatively consistent | Behavior is relatively inconsistent |
| Limited role of moral awareness | Heightened role of moral awareness |

dependent on what else she has decided and experienced recently, and how that affected her self-concept. If she started her day with a steady-state positive ethical self-view, she would have little moral self-threat, and thus be more likely to engage in a more aggressive forecasting decision, consistent with self-enhancement. Importantly, there is little moral awareness that flows through the self-enhancement processes, so she approaches the forecast as purely a business decision, not an ethical one. These processes are heavily automatic, and her ethical decision-making is more unconscious than conscious. In fact, in self-enhancement, her behavior will trend towards more and more unethical, with the forecast decision being one in a gradual progression.

On the other hand, if something had triggered her self-threat, such as being reminded that past practices of incorporating a cushion in projections are unacceptable, it might shift her towards self-protection. She might be in a state where her ethical self-view is under threat, sending her into self-protection. In this set of processes, moral awareness kicks in, and she will view the forecast as an ethical decision, making her more likely to report the cancelled order and less likely to report the hopeful future sale in her figures.

In a non-dynamic model, the sales manager's forecast would not vary as a function of her self-view and her self-view would not vary based on self-threats she had experienced prior to the forecast decision. She would make the same forecast, regardless of the various possible self-threats. Our model posits ethical behavior emerges from dynamic processing which means that the same individual can face the same decision multiple times, but behave differently on different occasions. This variability occurs because current and future behaviors are partially shaped by recent past behaviors, which will vary from occasion to occasion in their impact on self-view and degree of self-threat. Thus, the extent to which an individual remains in a particular mindset will hinge on whether her self-view has shifted. Given the dynamic nature of our model, we would anticipate that one's self-view would be in continual and dynamic flux. Ethical behavior is, thus, dynamic when what happens now is affected by what happened before.

3.4. How the model reconciles puzzles in the literature

3.4.1. Puzzle #1: When do we see the ethical implications of decisions and when don't we?

Moral awareness exists when an individual interprets a situation as containing ethical issues or as relevant to moral principles (Butterfield et al., 2000; Ferrell & Gresham, 1985; Rest, 1986; Reynolds, 2006b). Moral awareness is a function of both dispositional (moral attentiveness, Reynolds, 2008; ethical predispositions, Reynolds, 2006b; moral identity, Aquino & Reed, 2002) and situational (Treviño, 1986) variables.

Our model helps resolve when these situational variables lead to actual moral awareness. For example, studies have shown that use of honor codes signed before an ethical decision (Shu & Gino, 2012), a reminder of the Ten Commandments (Shu et al., 2011), the framing of a

decision (Tenbrunsel & Messick, 1999), or the physical environment (e.g. Liljenquist, Zhong, & Galinsky, 2010; Zhong & House, 2012) can heighten moral awareness. And, the reverse is also seen in the literature. Mazar et al. (2008) demonstrate that low moral awareness reduces the self-threat, which according to our model, leads to self-enhancement which permits a greater latitude in behavior.

In the context of bounded ethicality, we can interpret these manipulations as influencing self-threat assessment. By making one's ethical self-view or the potential for ethical self-threat more salient, these manipulations increase the likelihood that the self-threat assessment will generate a meaningful self-threat and lead to a self-protection process. In Fig. 1, self-enhancement leads directly to self-threat assessment, rather than to heightened moral awareness, while in contrast, self-protection will heighten moral awareness before self-threat assessment occurs. Certainly, other dispositional or situational factors may generate moral awareness during self-enhancement, but the previous behavior within the self-enhancement cycle will not contribute to this moral awareness.

This lack of awareness in the self-enhancement process is captured well in a model developed by Tenbrunsel and Smith-Crowe (2008), which distinguishes between decisions made with moral awareness and those made without moral awareness. In the “no moral awareness” condition, amoral decision-making processes lead to both ethical and unethical decisions. This kind of amoral decision making is what Tenbrunsel and Smith-Crowe (2008) refer to as “unintended” (in Table 2 and Fig. 1) and they speculate that the subset of decisions characterized by “unintended unethicality” may be the domain captured by bounded ethicality while encouraging future research to test this characterization.

We offer that our model of bounded ethicality incorporates both the intended and unintended groupings of the Tenbrunsel and Smith-Crowe (2008) typology, rather than just the unintended subset specified by the authors. That is, self-enhancement (with its limited role of moral awareness) aligns with the amoral decision making process, and self-protection (with its heightened role of moral awareness) aligns with the moral decision making process. In both self-enhancement and self-protection – as in both amoral and moral decision-making – the decision maker's behavior may be ethical or unethical. While moral awareness is neither a prerequisite for nor a guarantee of an ethical decision, moral awareness does shift individuals towards decision-making in which ethical decisions are more likely (Tenbrunsel & Smith-Crowe, 2008). We posit that the heightened role of moral awareness in self-protection contributes to the tendency for behavior to cycle towards greater ethicality in self-protection (see Table 2).

Because moral awareness, by definition, leads one to consider a situation through an ethical lens, it also increases the possibility of a self-threat becoming visible. That is, when we compare our self-view to the tolerance level under conditions of moral awareness, we are benefiting less from the rosier perspective that self-enhancement allows. Moral awareness makes the ethical

criteria for evaluation salient, and thus, leads to a potentially less generous evaluation of self-threat. Thus, moral awareness influences the degree of sensitivity to ethical self-threats and thus, makes it more likely that self-protection will be needed. We posit that moral awareness plays a relatively limited role in the self-enhancement process and a relatively heightened role in the self-protection process.

As Fig. 1 illustrates, these processes continue automatically, cyclically, and dynamically, shaping our ethical behavior. Moral awareness allows us to see more clearly the ethicality of our decisions, and reveals itself more fully when self-threat is present.

3.4.2. Puzzle #2: Does good behavior generate more good behavior, or license bad behavior?

This puzzle is best understood by reverse engineering what occurs leading up to the behavior. In the language of bounded ethicality, the results of the self-threat assessment shape the individual's likely next behaviors. Self-view is prone to situational influences. Because self-threat is shaped by the situation, one's past actions are continually feeding into what Monin and Jordan (2009) describe as one's "working level" response to the question "How moral am I?" and shaping one's present and future actions; this is the process we describe as self-threat assessment. This moment-to-moment assessment is not conscious or effortful; in our model, this is an automatic and ongoing process and the question might be more precisely described as "Is my ethical self-view in danger?" As long as the answer is no, the individual stays in self-enhancement.

The implication of this dynamic model is that the inputs into one's self-threat assessment are broad, fungible, and undulating. In other words, self-threat is volatile and ranges from minimal to meaningful. This perspective is also captured in self-concept maintenance theory (sometimes referred to as "fudge factor" theory, Mazar et al., 2008). The "fudge factor" represents the leeway one has in one's self-concept (what we refer to as self-view), which allows for some unethical behavior without damage to the self-concept.

The self-enhancement process has a distinctly cyclical nature. An individual is continually self-enhancing, viewing her ethical behavior as more ethical than it objectively is and viewing her unethical behavior to be less unethical than it objectively is. The very nature of self-enhancement is to positively distort how we view our own behavior, creating blind spots (Bazerman & Tenbrunsel, 2012). As a result, according to the bounded ethicality model, she is less likely to perceive a self-threat. As seen in Fig. 1, the lack of self-threat (following the "no" path in Fig. 1) sends her back into the self-enhancement cycle. The consequence of this cyclical tendency is a critical feature of our model, and a potentially slippery slope.

As one continually self-enhances, one is also continually cycling through the process of not seeing gaps in one's own behavior and overvaluing one's ethical behavior. Thus, our model would suggest that one is continually overestimating one's moral credits (Miller & Effron, 2010) and the slack in the moral tension (Zhong et al., 2009), pushing

us with each cycle of self-enhancement to move towards more unethical behavior. The momentum during self-enhancement cycles is towards the inner loop (solid line) of unethical behavior and away from the outer loop (dotted line) of ethical behavior, as seen in Fig. 1. Another way to think about this pattern is that the self-enhancement trends away from primary control (being ethical) and towards secondary control (feeling ethical).

The result is behavior that is sometimes consistent with past behavior, such as when I perform multiple good deeds throughout the day, versus behavior that is sometimes in contrast with past behavior, such as when I perform a good deed in the afternoon, followed by a bad deed soon after that. This distinction between consistency (good leads to good, bad leads to bad) and contrast (good leads to bad, or bad leads to good) patterns of behavior is well-captured and analyzed by Miller and Effron (2010). They highlight that one's "identity" may be a key determinant of which pattern unfolds, and in our model, we push this notion even further by describing the dynamic and cyclical process of continually assessing the self-threat levels. Our model predicts that when ethical self-threat is high, there will be a more urgent psychological need to self-protect, leading to a tendency to good behavior. If bad behavior had led to the ethical self-threat, then this is where a contrast pattern will emerge (unethical to ethical behavior). If the ethical self-threat has emerged, despite a good deed, then a consistency pattern will emerge (ethical to ethical behavior). When ethical self-threat is low, our model shows that self-enhancement will pull us towards the bad, even if we began with a good deed (ethical to unethical). In other words, our model details the mechanism behind these otherwise conflicting predictions about consistency versus contrast patterns.

This pattern creates the conditions under which ordinary and small unethical behaviors can lead to far less ordinary and far more significant unethical behaviors over time, until there is a distinct self-threat which interrupts this process. This slippery slope phenomenon (Gino & Bazerman, 2009; Gino et al., 2009) is partially due to bounded awareness (Bazerman & Chugh, 2005), the tendency to overlook important and easily available information. Both slippery slope processes – within one's own behavior during self-enhancement and when judging others' behavior – carry elements of an integrated motivational and cognitive perspective. The blind spots in both behaviors and judgments are not purely cognitive "misses" in which a perceptual error occurs, but are systematically oriented towards particular types of misses, motivated to favor a particular self-view. Similarly, self-enhancement draws people towards the goal of being ethical (as opposed to not being unethical), which is similar to a promotion (as opposed to prevention) focus (Higgins, 1997, 1998). Promotion focus has been shown to increase unethical behavior (Gino & Margolis, 2011).

This tendency towards more unethical behavior over time is due to a process in which self-enhancement leaves one at ease with the previous act in the end, even if it might have originally seemed like a 'gray area'. During self-enhancement, slight indiscretions are easily incorporated into one's self-view without consequence (Alicke &

Sedikides, 2009). What was once gray area, is no longer gray area, and the cycle continues from there, until the self-threat assessment process generates a sufficiently large self-threat to interrupt self-enhancement and initiate self-protection. Until this occurs, during self-enhancement, it is likely that individuals tend towards increasingly more unethical behavior.

The effects of the cyclical tendency are also seen during the self-protection cycles, but with reverse consequences. One might imagine individuals sliding down the slippery slope of self-enhancement until they hit the bump of self-threat, shifting them towards self-protection. Self-protection, by definition, is a more episodic and less automatic process that is prompted by a self-threat, which instigates a drive to restore one's self-view. The self-protection cycle (following the "yes" path in Fig. 1) is therefore designed to either reduce the self-threat through secondary control (unethical behavior) or primary control (ethical behavior). Regardless of whether the behavior is ethical or unethical, one's moral awareness will be heightened due to the ethical self-threat. That is, self-protection leads to a boost in moral awareness, and moral awareness leads to more ethical behavior (Shu et al., 2011), and thus, subsequent cycles of self-protection will lean more and more towards ethical behavior, until the cycle through self-threat assessment reveals that the ethical self-threat has been removed. At this point, the objective of self-protection has been achieved – the self-threat has been reduced – and the individual can resume the continuous process of self-enhancement (following the "no" path in Fig. 1). Thus, the momentum during the self-protection cycles is towards the outer loop (solid line) of ethical behavior and away from the inner loop (dotted line) of unethical behavior. As a result, it is likely that during self-protection, individuals tend towards increasingly more ethical behavior.

Bounded ethicality provides specificity regarding when current behavior is more or less consistent with previous behavior (see Table 2). When on the slippery slope of self-enhancement, individuals are more likely to gradually move towards unethical behavior and secondary control strategies, and thus, there is relative consistency in behavior. That is, unethical behaviors are likely to follow unethical behaviors in self-enhancement. Self-protection, on the other hand, is characterized by a relative lack of consistency in behaviors. Individuals are trying to restore their self-view and thus, when trending towards a primary control strategy, will behave in contrast to their previous actions, resulting in movement in the direction of ethical behavior.

Interestingly, this pattern is also seen in work on psychological license (Miller & Effron, 2010), which distinguishes between moral credits and moral credentials. Moral credentials are also characterized by relative consistency (like self-enhancement) whereas moral credits are characterized by relative inconsistency (like self-protection). The original moral credentials work demonstrated that individuals were more likely to express views that might be seen as prejudiced if they had previously "established their credentials as nonprejudiced persons" (Monin & Miller, 2001). Once these credentials are in place, individuals reconstrue bad behaviors into not-bad, or even

good, behaviors, a process akin to self-enhancement. Thus, the moral credentials model aligns with self-enhancement as it reconstrues behaviors are no longer potentially self-threatening (Monin & Miller, 2001; Miller & Effron, 2010). In contrast, in the moral credits model, "bad" behavior is not reconstrued, but is still seen as bad behavior which consumes the moral credits generated by good behavior. Here, individuals are more likely to behave inconsistently with previous behavior, and we propose that this is parallel to the self-protection cycle in which a bad behavior – seen as a bad behavior – is a self-threat. That is, it is more likely that individuals will show greater consistency in their behavior during self-enhancement than during self-protection.

The reader may notice that we are positing a series of asymmetric patterns in our model. Next, we will focus on another asymmetry between self-enhancement and self-protection. Specifically, we will focus on the role of moral awareness in the process of ethical decision-making, and how that role differs in self-protection versus self-enhancement.

3.4.3. Puzzle #3: Why are we sometimes blind to our own transgressions?

When one is able to engage in unethical behavior without feeling unethical, one might actually be blind to this disconnect. Bounded ethicality focuses less on how ethical one's behavior actually is and more on the psychological processes which enable individuals to remain confident of their own ethicality, despite ethical lapses.

Secondary control mechanisms are particularly useful in this regard. For example, after committing an unethical act, individuals are more likely to forget moral rules prohibiting that act; "forgetting" the moral rules is a secondary control mechanism that eliminates the self-threat that occurs when one's behavior violates the moral rules (Shu & Gino, 2012).

Another example of secondary control is found in the work on necessary evils (Margolis & Molinsky, 2008; Molinsky & Margolis, 2005). A necessary evil is "a work-related task that requires a person to cause physical, emotional, or material harm to another human being in order to advance a perceived greater good" (Molinsky & Margolis, 2005), such as conducting layoffs. For the manager conducting the layoffs, he or she faces a self-threat as his perceived level of ethicality may fall short of his tolerance level, thus threatening his self-view as someone who does not harm others. A manager in this position must address this self-threat in order to conduct the duties required by his role. A primary control approach would be to not conduct the layoff. A secondary control approach is to frame the task as necessarily due to changes in the global marketplace, and not due to managerial failures. With this reframing, the manager can maintain his or her positive ethical self-view, while still conducting the layoff. In subsequent layoffs, he may experience reduced moral awareness, making it more likely that the self-threat-triggered self-protection process will be replaced by the more comforting self-enhancement process.

Interestingly, secondary control mechanisms illustrate that self-enhancement is not a merely superficial process that operates independently of one's true views. When individuals self-enhance their self-views, they may come to actually believe the enhanced version of "deep self-deception" (Fingarette, 2000), so much so that they make consequential forecasting errors about their future performance based on the self-enhanced version of the past and present (Chance, Norton, Gino, & Ariely, 2011). In contrast, self-enhancement can also manifest as "light self-deception" (Fingarette, 2000), akin to not opening mail that contains bills requiring payment (Greenwald, 1980).

Self-enhancement plays a particularly pernicious role in generating blind spots. The processes of self-enhancement create a boost in one's self-view, so that one's unethical acts appear less unethical and one's ethical acts appear more ethical. Because self-enhancement processes are automatic, one does not realize that this insidious process has occurred.

Our blind spots contribute to Mazar, Amir, and Ariely's (2008) "band of acceptable dishonesty". Operating in this band allows individuals to capitalize on the gains that unethical behavior might bring, but not to the maximal degree possible. Small instances of unethical behavior do not pose self-threats as they are easily finessed through self-enhancement's secondary control mechanisms.

Similarly, in the moral credits model, which includes compensatory ethics (Zhong et al., 2010) and moral self-regulation (Zhong et al., 2009), "moral tension" is akin to the tension of a rubber band. Slack occurs when an individual has "moral credits" which allows individuals to relax their moral behavior, which aligns with our description of self-enhancement's capacity to handle ethical lapses through secondary control mechanisms, while a tightened band suggested that the individual's self-perception is highly misaligned with their self-view, thus generating a meaningful self-threat. In our model, self-enhancement is continually generating blind spots, which prevent individuals from perceiving ethical self-threats.

Blind spots are also aligned with Sonenshein's (2007) sensemaking-intuition model. Sensemaking (Weick, 1995) is the process by which people give meaning to experience, especially under conditions of uncertainty (plausible interpretations for a situation are not clear) and equivocality (multiple interpretations exist for a situation). Sonenshein (2007: 1029) argues that people's "motivational drives" are part of the sensemaking process in ethical decision-making via "what they want to see". Specifically, when searching for a plausible interpretation amidst multiple interpretations, individuals will favor the interpretation that is consistent with the positive ethical self-view to which they aspire. Thus, the process of sensemaking is well-captured in self-enhancement and its potential for blind spots.

In contrast, self-protection is characterized by a desire to not be bad (rather than to be good), which creates less of a motivational pull towards a positive self-view and more of a motivational pull away from a negative self-view. This desire for self-protection is stronger than the desire for self-enhancement (Alicke & Sedikides, 2009) and has the potential to generate a heightened moral awareness. The

role of subconscious primes in this heightened moral awareness is useful to consider. Moral reminders, even outside of awareness, lead to more ethical behavior through the mechanism of increasing one's tendency to consider situations through an ethical lens (Welsh & Ordóñez, 2014). So, the inclination is towards primary control, rather than secondary control. Empirically, this mechanism provides support for the role of moral awareness in our model, and broadens our understanding of what is meant by moral awareness.

In summary, the stronger motivation to not be bad creates a heightened attention towards self-threat assessment (the comparison between perceived and tolerance levels in one's ethical self-view). Thus, moral awareness is higher, self-threat is more accurately perceived and blind spots are less frequent during self-protection than in self-enhancement. In contrast, individuals are more prone to blind spots during self-enhancement than during self-protection, when moral awareness is muted and our decisions and behaviors are processed with a positive spin. The most important consequence of this difference between self-enhancement and self-protection is that we are less likely to not see our own ethical shortcomings when self-threat is high, because self-protection is less likely to blind us to those shortcomings. We will move towards more ethical behavior, as a result.

4. An agenda for future research

We offer an overarching model of ethical decision-making in which we specify the ways in which ethicality is systematically bounded, or "bounded ethicality". Our model describes "the systematic and ordinary psychological processes of enhancing and protecting our ethical self-view which automatically, dynamically, and cyclically influence the ethicality of decision-making". Our model does not explain all ethical decision-making, but brings important insights to everyday behaviors in which individuals operate in organizational contexts and make decisions affecting others.

We offer two central claims. First, we argue that the self (rather than self-interest) is the central driver of ethical decision-making, rooting our arguments in several key constructs from the self-literature. Second, we specify the role which automaticity plays in ethical decision-making and the conditions under which behavior is more or less automatic.

The model includes three key processes: self-threat assessment, self-protection, and self-enhancement. Self-threat assessment is the process of determining if an individual's self-view is threatened and determines if self-enhancement or self-protection will follow. In the absence of self-threat, the continual process of self-enhancement maintains its orientation towards the goal of a positive ethical self-view; in the presence of a self-threat, self-enhancement yields to the more episodic process of self-protection. Relative to self-protection, self-enhancement generates relatively more automatic processing, more blind spots, more unethical behavior, more consistent behavior, and more limited moral awareness. These distinct, asymmetric patterns integrate and synthesize

many of the key models and findings in recent behavioral ethics research into a single, overarching model of ethical decision-making.

Our model has useful implications for research about ethics in organizations and research about the self.

4.1. Ethics research

Recent pushes to take ethics research from the lab and build in the complexity and realities of organizations (Brief & Smith-Crowe, 2016; Zhang, Gino, & Bazerman, 2014) are particularly compelling. We hope our model offers an anchor for such work, perhaps through field experiments which examine how ethical self-threat influences subsequent decisions. We also provide a foundation for greater study of the role of automaticity. Many ethics scholars have increasingly alluded to the possibility that ethical behavior and judgments may not be fully deliberative (Haidt, 2002; Reynolds et al., 2010; Sonenshein, 2007) and we build on these speculations with theorizing about the specific nature of the automatic processing, providing scaffolding for future work in this area. This specificity is important to the development of prescriptive approaches as it will allow us to match the automaticity of the intervention to the automaticity of the underlying process, which is an important feature of designing effective interventions (Milkman, Chugh, & Bazerman, 2009).

With this understanding of ethical behavior in which everyone is prone to ethical lapses, organizations are better positioned to influence the behaviors of employees through the design of effective ethical interventions. In an organization, behavior is shaped both by formal structure and informal culture (Nadler & Tushman, 1980). Through both formal structures and informal cultures, organizations are continually heightening or diminishing ethical self-threat, and this is an important direction for future research. For example, an expense reimbursement system that requires extensive documentation and allows for no flexibility on lost receipts is likely to heighten self-threat, as it sends a signal that the employee is likely to cheat unless constrained from doing so. Thus, the presence or absence of ethical self-threat is not only a function of one's past behaviors, but also of one's environment.

Specifically, while we have shown that the presence and absence of ethical self-threat contributes to our ethical failures and successes, the relationship is not a straightforward one in which more ethical self-threat is clearly positive (or negative). The magnitude of self-threat determines whether the self-enhancement or self-protection processes are active, and depending on which process is active, different patterns of ethical behavior unfold. However, the pattern of behavior that tends towards ethical behavior emerges during self-protection, suggesting that sustained self-enhancement, without some episodic self-threat, is ethically treacherous for individuals and organizations and an important area for study.

This pattern suggests that ethical behavior within organizations would benefit from two types of "nudges" (Thaler & Sunstein, 2008). One set of nudges would

heighten self-threat so as to nudge people out of self-enhancement and towards self-protection. Once in self-protection, people are more likely to move towards ethical behavior. Another set of nudges would heighten moral awareness, similar to the effect that occurs during self-protection, and thus, generate momentum towards ethical behavior. The work of Welsh and Ordóñez (2014) demonstrates the effects of subconscious ethical and unethical priming in reducing dishonesty when participants are unmonitored, suggesting that the nudges can operate subtly by creating moral awareness, even in the absence of an ethical self-threat. Future research might explore these nudges, both in isolation as well as in combination with other interventions shown to improve ethical behavior. For example, recent work shows the benefits of encouraging people to think about what they "could" do (rather than what they "should" do) (Zhang, Gino, & Margolis, 2014) and of asking for a signature certifying honesty before decisions are made rather than after (Shu, Mazar, Gino, Ariely, & Bazerman, 2012). The common thread of this work is that the rules dictating what is or is not ethical have not changed nor has the incentive against or penalty for unethical behavior changed. Rather, a small nudge has heightened the decision-maker's moral awareness and increased the probability of a decision made through the self-protection cycle.

Further research on ethical learning might leverage the bounded ethicality model and offer rich possibilities for research in organizations. Ethical learning occurs when individuals care about being ethical, understand that they are not always as ethical as they wish to be, and believe that they can grow into being more ethical (Chugh & Kern, 2016). According to this model, when this type of ethical learning occurs in an organization or team which creates a shared belief that the team is a safe context for ethical learning, actual improvements in ethical behavior become possible. Exploring how nudges can move individuals, teams, and organizations in this direction is an exciting direction for future research.

The central role of self-view also highlights the need for researchers to distinguish between what we know about people's ethical behavior and what we know about how people judge other people's ethical behavior. Bounded ethicality implies that the psychological processes underlying one's own behavior versus one's judgments of others are distinct because self-view plays a more central role in the former than the latter. Bounded ethicality, and this proposed model focus on one's own ethical behavior. Work on moral hypocrisy (Batson & Collins, 2011; Kreps & Monin, 2011; Lammers, Stapel, & Galinsky, 2010; Valdesolo & DeSteno, 2007) also highlights this disconnect between behavior and judgments. While work on moral exemplars suggests that our self-view distorts how we view the behaviors of others (Monin, Sawyer, & Marquez, 2008), we suspect that judgments of others are not as determinant of one's self-view as one's own behavior. Thus, we speculate that these judgments are less influenced by self-enhancement and self-protection and that an important boundary condition of our model may be that it applies to the self more than to others. This

potential boundary condition points to the value of closer future study of how these processes are similar or different.

Haidt's social intuitionist model (2001) is an excellent example of a model that aligns well with bounded ethicality but with different boundary conditions. In this model, moral rules are applied through automatic processes, leading to reflexive judgments for which individuals build reasoned support after the judgment has already formed spontaneously. Like bounded ethicality, the social intuitionist model highlights the role of automaticity, but unlike bounded ethicality, the role of the self is greatly diminished. The social intuitionist model is focused on how we reach judgments of others' ethical behavior, while bounded ethicality is focused on our own ethical behavior.

4.2. Self research

Our model also shows the prescriptive value of important constructs from the self literature and demonstrates what is necessary for these descriptive constructs to have real-world implications. Well-studied processes, such as self-enhancement and self-protection, are powerful tools in the study of behavior important to both organizations and societies. One useful future direction is to explore when processes such as self-verification (Swann & Read, 1981; processes that are oriented towards aligning how others view an individual with how the individual views herself) surpass processes like self-enhancement and self-protection in importance in ethical decision-making. Future research might also generate clearer insight into when self-view trumps self-interest, and vice versa.

This model of bounded ethicality also provides a peek into the need to be precise in work which focuses on "self as subject" (I-self) versus work that focuses on "self as object" (me-self) (Brockner & Wiesenfeld, forthcoming; Tangney & Leary, 2012). In our model, the self-view is an example of my view of myself, analogous to how I would form a view of an external object; this is the "self as object" use of the self. In contrast, we are not using the "self as subject" perspective, which is better illustrated by work on self-regulation (Gino et al., 2011; Mead et al., 2009). In this research, one is not forming a view of oneself, but experiencing oneself. In future work, self researchers and ethics researchers will benefit from teasing apart these two very different applications of the self.

5. Conclusion

In closing, our model of bounded ethicality integrates much of the recent research in the active and thriving study of behavioral ethics. Our hope is that it provides scholars and practitioners alike a cohesive platform for understanding ethical failures and designing ethical interventions. In particular, we seize the opportunity to leverage the insights about the self and automaticity in order to deepen our knowledge of bounded ethicality and look forward to the future research building on bounded ethicality 2.0.

References

- Alicke, M. D., & Sedikides, C. (2009). Self-enhancement and self-protection: What they are and what they do. *European Review of Social Psychology*, 20(1), 1–48.
- Allison, S. T., & Messick, D. M. (1990). Social decision heuristics in the use of shared resources. *Journal of Behavioral Decision Making*, 3(3), 195–204.
- Aquino, K., & Reed, A. (2002). The self-importance of moral identity. *Journal of Personality and Social Psychology*, 83(6), 1423–1440.
- Babcock, L., & Loewenstein, G. (1997). Explaining bargaining impasse: The role of self-serving biases. *Journal of Economic Perspectives*, 11, 109–126.
- Banaji, M. R., & Bhaskar, R. (2000). Implicit stereotypes and memory: The bounded rationality of social beliefs. *Memory, Brain, and Belief*, 139–175.
- Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice Hall.
- Batson, C. D., & Collins, E. C. (2011). Moral hypocrisy: A self-enhancement/self-protection motive in the moral domain. In *Handbook of self-enhancement and self-protection* (pp. 92–111). New York: Guilford Press.
- Baumeister, R. F., Smart, L., & Boden, J. M. (1996). Relation of threatened egotism to violence and aggression: The dark side of high self-esteem. *Psychological Review*, 103(1), 5–33.
- Baumeister, R. F., Vohs, K. D., & Tice, D. M. (2007). The strength model of self-control. *Current Directions in Psychological Science*, 16(6), 351–355.
- Baumhart, R. (1968). *Ethics in business*. New York: Holt, Rinehart and Winston.
- Bazerman, M. H., & Chugh, D. (2005). Bounded awareness: Focusing problems in negotiation. In L. Thompson (Ed.), *Frontiers of social psychology: Negotiation* (pp. 7–26). College Park, MD: Psychology Press.
- Bazerman, M. H., & Tenbrunsel, A. E. (2012). *Blind spots: Why we fail to do what's right and what to do about it*. Princeton, NJ: Princeton University Press.
- Blasi, A. (2004). Moral functioning: Moral understanding and personality. In D. Narvaez & D. K. Lapsley (Eds.), *Moral development, self, and identity* (pp. 335–347). New York: Cambridge University Press.
- Brief, A. P., & Smith-Crowe, K. (2016). Why organizations matter. In A. G. Miller (Ed.), *The social psychology of good and evil* (2nd ed.). New York: Guilford Press.
- Brockner, J., & Wiesenfeld, B. (2016). Self-as-subject and self-as-object in the workplace. *Organizational Behavior and Human Decision Processes* (forthcoming).
- Bryan, C. J., Adams, G. S., & Monin, B. (2013). When cheating would make you a cheater: Implicating the self prevents unethical behavior. *Journal of Experimental Psychology: General*, 142(4), 1001–1005.
- Bushman, B. J., Moeller, S. J., & Crocker, J. (2011). Sweets, sex, or self-esteem? Comparing the value of self-esteem boosts with other pleasant rewards. *Journal of Personality*, 79(5), 993–1012.
- Butterfield, K. D., Treviño, L. K., & Weaver, G. R. (2000). Moral awareness in business organizations: Influences of issue-related and social context factors. *Human Relations*, 53(7), 981–1018.
- Campbell, W. K., & Sedikides, C. (1999). Self-threat magnifies the self-serving bias: A meta-analytic integration. *Review of General Psychology*, 3(1), 24–43.
- Caruso, E., Epley, N., & Bazerman, M. H. (2006). The costs and benefits of undoing egocentric responsibility assessments in groups. *Journal of Personality and Social Psychology*, 91(5), 857–871.
- Chance, Z., Norton, M. I., Gino, F., & Ariely, D. (2011). Temporal view of the costs and benefits of self-deception. *Proceedings of the National Academy of Sciences of the United States of America*, 108(Suppl. 3), 15655–15659.
- Chugh, D., Banaji, M., & Bazerman, M. (2005). Bounded ethicality as a psychological barrier to recognizing conflicts of interest. In D. Moore, D. Cain, G. Loewenstein, & M. Bazerman (Eds.), *Conflicts of interest: Challenges and solutions in business, law, medicine, and public policy*. New York: Cambridge University Press.
- Chugh, D., & Kern, M. C. (2016). Becoming as ethical as we think we are: The ethical learner at work. In D. A. Palmer, R. Greenwood, & K. Smith-Crowe (Eds.), *Organizational wrongdoing*. New York: Cambridge University Press.
- Detert, J. R., Treviño, L. K., & Sweitzer, V. L. (2008). Moral disengagement in ethical decision making: A study of antecedents and outcomes. *Journal of Applied Psychology*, 93(2), 374–391.
- Donaldson, T., & Dunfee, T. W. (1994). Toward a unified conception of business ethics: Integrative social contracts theory. *Academy of Management Review*, 19(2), 252–284.
- Effron, D. A., & Monin, B. (2010). Letting people off the hook: When do good deeds excuse transgressions? *Personality and Social Psychology Bulletin*, 36(12), 1618–1634.
- Einhorn, H. J., & Hogarth, R. M. (1978). Confidence in judgment: Persistence of the illusion of validity. *Psychological Review*, 85(5), 395–416.

- Epley, N., & Caruso, E. M. (2004). Egocentric ethics. *Social Justice Research, 17*(2), 171–187.
- Fehr, E., & Fischbacher, U. (2003). The nature of human altruism. *Nature, 425*(6960), 785–791.
- Ferrell, O., & Gresham, L. G. (1985). A contingency framework for understanding ethical decision making in marketing. *Journal of Marketing, 87*–96.
- Fingarette, H. (2000). *Self-deception: With a new chapter*. Berkeley, CA: University of California Press.
- Gilbert, D. T., & Hixon, J. G. (1991). The trouble of thinking: Activation and application of stereotypic beliefs. *Journal of Personality and Social Psychology, 60*(4), 509–517.
- Gino, F., & Bazerman, M. H. (2009). When misconduct goes unnoticed: The acceptability of gradual erosion in others' unethical behavior. *Journal of Experimental Social Psychology, 45*(4), 708–719.
- Gino, F., Moore, D. A., & Bazerman, M. H. (2009). See no evil: When we overlook other people's unethical behavior. In R. M. Kramer, A. E. Tenbrunsel, & M. H. Bazerman (Eds.), *Social decision making: Social dilemmas, social values, and ethical judgments* (pp. 241–263). New York: Psychology Press.
- Gino, F., & Margolis, J. D. (2011). Bringing ethics into focus: How regulatory focus and risk preferences influence (un)ethical behavior. *Organizational Behavior and Human Decision Processes, 115*(2), 145–156.
- Gino, F., Schweitzer, M. E., Mead, N. L., & Ariely, D. (2011). Unable to resist temptation: How self-control depletion promotes unethical behavior. *Organizational Behavior and Human Decision Processes, 115*(2), 191–203.
- Gneezy, U., Imas, A., & Madarász, K. (2014). Conscience accounting: Emotion dynamics and social behavior. *Management Science, 60*(11), 2645–2658.
- Grant, A. M. (2007). Relational job design and the motivation to make a prosocial difference. *Academy of Management Review, 32*(2), 393–417.
- Greenwald, A. G. (1980). The totalitarian ego: Fabrication and revision of personal history. *American Psychologist, 35*(7), 603–618.
- Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review, 102*(1), 4–27.
- Haidt, J. (2001). The emotional dog and its rational tail: A social intuitionist approach to moral judgment. *Psychological Review, 108*(4), 814–834.
- Haidt, J. (2002). Dialogue between my head and my heart: Affective influences on moral judgment. *Psychological Inquiry, 13*, 54–56.
- Haidt, J. (2008). Morality. *Perspectives on Psychological Science, 3*, 65–72.
- Higgins, E. T. (1987). Self-discrepancy: A theory relating self and affect. *Psychological Review, 94*(3), 319–340.
- Higgins, E. T. (1997). Beyond pleasure and pain. *American Psychologist, 52*(12), 1280–1300.
- Higgins, E. T. (1998). Promotion and prevention: Regulatory focus as a motivational principle. *Advances in Experimental Social Psychology, 30*, 1–46.
- Jolls, C., Sunstein, C. R., & Thaler, R. (1998). A behavioral approach to law and economics. *Stanford Law Review, 50*(5), 1471–1550.
- Jones, T. M. (1991). Ethical decision making by individuals in organizations: An issue-contingent model. *Academy of Management Review, 16*(2), 366–395.
- Jordan, J., Mullen, E., & Murnighan, J. K. (2011). Striving for the moral self: The effects of recalling past moral actions on future moral behavior. *Personality and Social Psychology Bulletin, 37*(5), 701–713.
- Kern, M. C., & Chugh, D. (2009). Bounded ethicality: The perils of loss framing. *Psychological Science, 20*(3), 378–384.
- Kohlberg, L. W. (1969). *Stage and sequence: The cognitive development approach to socialization*. New York: Rand McNally.
- Kreps, T. A., & Monin, B. (2011). Doing well by doing good? Ambivalent moral framing in organizations. *Research in Organizational Behavior, 31*, 99–123.
- Krusemark, E. A., Campbell, W. K., & Clementz, B. A. (2008). Attributions, deception, and event related potentials: An investigation of the self-serving bias. *Psychophysiology, 45*(4), 511–515.
- Lammers, J., Stapel, D. A., & Galinsky, A. D. (2010). Power increases hypocrisy moralizing in reasoning, immorality in behavior. *Psychological Science, 21*(5), 737–744.
- Leary, M. R., Terry, M. L., Allen, A. B., & Tate, E. B. (2009). The concept of ego threat in social and personality psychology: Is ego threat a viable scientific construct? *Personality and Social Psychology Review, 13*(3), 151–164.
- Liljenquist, K., Zhong, C. B., & Galinsky, A. D. (2010). The smell of virtue clean scents promote reciprocity and charity. *Psychological Science, 21*(3), 381–383.
- Margolis, J. D. (2009). The responsibility gap. *Hedgehog Review, 11*(2), 41–53.
- Margolis, J. D., & Molinsky, A. (2008). Navigating the bind of necessary evils: Psychological engagement and the production of interpersonally sensitive behavior. *Academy of Management Journal, 51*(5), 847–872.
- Mazar, N., Amir, O., & Ariely, D. (2008). The dishonesty of honest people: A theory of self-concept maintenance. *Journal of Marketing Research, 45*(6), 633–644.
- Mead, N. L., Baumeister, R. F., Gino, F., Schweitzer, M. E., & Ariely, D. (2009). Too tired to tell the truth: Self-control resource depletion and dishonesty. *Journal of Experimental Social Psychology, 45*(3), 594–597.
- Messick, D. M., & Bazerman, M. H. (1996). Ethical leadership and the psychology of decision making. *Sloan Management Review, 37*, 9–22.
- Milkman, K. L., Chugh, D., & Bazerman, M. H. (2009). How can decision making be improved? *Perspectives on Psychological Science, 4*(4), 379–383.
- Miller, D. T. (1999). The norm of self-interest. *American Psychologist, 54*, 1–8.
- Miller, D. T., & Efron, D. A. (2010). Psychological license: When it is needed and how it functions. *Advances in Experimental Social Psychology, 43*, 115–155.
- Molinsky, A., & Margolis, J. (2005). Necessary evils and interpersonal sensitivity in organizations. *Academy of Management Review, 30*(2), 245–268.
- Monin, B., & Jordan, A. H. (2009). The dynamic moral self: A social psychological perspective. In D. Narvaez & D. K. Lapsley (Eds.), *Personality, identity, and character: Explorations in moral psychology* (pp. 341–354). New York: Cambridge University Press.
- Monin, B., & Miller, D. T. (2001). Moral credentials and the expression of prejudice. *Journal of Personality and Social Psychology, 81*(1), 33–43.
- Monin, B., Sawyer, P. J., & Marquez, M. J. (2008). The rejection of moral rebels: Resenting those who do the right thing. *Journal of Personality and Social Psychology, 95*(1), 76–93.
- Moore, D. A., & Loewenstein, G. (2004). Self-interest, automaticity, and the psychology of conflict of interest. *Social Justice Research, 17*(2), 189–202.
- Moore, D. A., Cain, D. M., Loewenstein, G., & Bazerman, M. H. (2005). *Conflicts of interest*. Cambridge, UK: Cambridge University Press.
- Murnighan, J. K., Canteloni, D. A., & Elyashiv, T. (2001). Bounded personal ethics and the tap dance of real estate agency. *Advances in Qualitative Organizational Research, 3*, 1–40.
- Nadler, D. A., & Tushman, M. L. (1980). A model for diagnosing organizational behavior. *Organizational Dynamics, 9*(2), 35–51.
- Nisan, M. (1991). The moral balance model: Theory and research extending our understanding of moral choice and deviation. In W. M. Kurtines & J. L. Gewirtz (Eds.), *Handbook of moral behavior and development* (pp. 213–249). Hillsdale, NJ: Erlbaum.
- Nuttin, J. M. (1985). Narcissism beyond Gestalt and awareness: The name letter effect. *European Journal of Social Psychology, 15*(3), 353–361.
- Paulhus, D. L., & Levitt, K. (1987). Desirable responding triggered by affect: Automatic egotism? *Journal of Personality and Social Psychology, 52*(2), 245–259.
- Pitesa, M., & Thau, S. (2013). Compliant sinners, obstinate saints: How power and self-focus determine the effectiveness of social influences in ethical decision making. *Academy of Management Journal, 56*(3), 635–658.
- Rest, J. R. (1986). *Moral development: Advances in research and theory*. New York: Praeger.
- Reynolds, S. J. (2006a). A neurocognitive model of the ethical decision-making process: Implications for study and practice. *Journal of Applied Psychology, 91*(4), 737–748.
- Reynolds, S. J. (2006b). Moral awareness and ethical predispositions: Investigating the role of individual differences in the recognition of moral issues. *Journal of Applied Psychology, 91*(1), 233–243.
- Reynolds, S. J. (2008). Moral attentiveness: Who pays attention to the moral aspects of life? *Journal of Applied Psychology, 93*(5), 1027–1041.
- Reynolds, S. J., Leavitt, K., & DeCelles, K. A. (2010). Automatic ethics: The effects of implicit assumptions and contextual cues on moral behavior. *Journal of Applied Psychology, 95*(4), 752–760.
- Rothbaum, F., Weisz, J. R., & Snyder, S. S. (1982). Changing the world and changing the self: A two-process model of perceived control. *Journal of Personality and Social Psychology, 42*(1), 5.
- Sedikides, C. (2009). On self-protection and self-enhancement regulation: The role of self-improvement and social norms. In J. P. Forgas, R. F. Baumeister, & D. M. Tice (Eds.), *The psychology of self-regulation: Cognitive, affective, and motivational processes* (pp. 73–92). New York: Psychology Press.
- Sedikides, C. (2012). Self-protection. In M. R. Leary, & J. P. Tangney (Eds.), *Handbook of self and identity* (2, pp. 327–353). The Guilford Press.
- Shu, L. L., & Gino, F. (2012). Sweeping dishonesty under the rug: How unethical actions lead to forgetting of moral rules. *Journal of Personality and Social Psychology, 102*(6), 1164–1177.
- Shu, L. L., Gino, F., & Bazerman, M. H. (2011). Dishonest deed, clear conscience: When cheating leads to moral disengagement and motivated forgetting. *Personality and Social Psychology Bulletin, 37*(3), 330–349.
- Shu, L. L., Mazar, N., Gino, F., Ariely, D., & Bazerman, M. H. (2012). Signing at the beginning makes ethics salient and decreases dishonest self-reports in comparison to signing at the end. *Proceedings of the National Academy of Sciences of the United States of America, 109*(38), 15197–15200.
- Simon, H. A. (1957). *Models of man: social and rational*. Oxford: Wiley.
- Sonenshein, S. (2007). The role of construction, intuition, and justification in responding to ethical issues at work: The sensemaking-intuition model. *Academy of Management Review, 32*(4), 1022–1040.

- Strohinger, N., & Nichols, S. (2014). The essential moral self. *Cognition*, *131*, 159–171.
- Sunstein, C. R. (2005). Moral heuristics. *Behavioral and Brain Sciences*, *28*(4), 531–541.
- Swann, W. B., Jr., & Read, S. J. (1981). Self-verification processes: How we sustain our self-conceptions. *Journal of Experimental Social Psychology*, *17*, 351–372.
- Swann, W. B., Jr. (1983). Self-verification: Bringing social reality into harmony with the self. *Social psychological perspectives on the self*, *2*, 33–66.
- Tangney, J. P., & Leary, M. R. (Eds.). (2012). *Handbook of self and identity*. New York: Guilford Press.
- Tenbrunsel, A. E. (1998). Misrepresentation and expectations of misrepresentation in an ethical dilemma: The role of incentives and temptation. *Academy of Management Journal*, *41*(3), 330–339.
- Tenbrunsel, A. E., & Messick, D. M. (1999). Sanctioning systems, decision frames, and cooperation. *Administrative Science Quarterly*, *44*(4), 684–707.
- Tenbrunsel, A. E., & Smith-Crowe, K. (2008). Ethical decision making: Where we've been and where we're going. *Academy of Management Annals*, *2*(1), 545–607.
- Tenbrunsel, A. E., Diekmann, K. A., Wade-Benzoni, K. A., & Bazerman, M. H. (2010). The ethical mirage: A temporal explanation as to why we are not as ethical as we think we are. *Research in Organizational Behavior*, *30*, 153–173.
- Thaler, R. H., & Sunstein, C. R. (2008). *Nudge: Improving decisions about health, wealth, and happiness*. London: Yale University Press.
- Treviño, L. K. (1986). Ethical decision making in organizations: A person-situation interactionist model. *Academy of Management Review*, *11*(3), 601–617.
- Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, *185*(4157), 1124–1131.
- Valdesolo, P., & DeSteno, D. (2007). Moral hypocrisy: Social groups and the flexibility of virtue. *Psychological Science*, *18*(8), 689–690.
- Victor, B., & Cullen, J. B. (1988). The organizational bases of ethical work climates. *Administrative Science Quarterly*, *33*, 101–125.
- Vohs, K. D., Baumeister, R. F., & Ciarocco, N. J. (2005). Self-regulation and self-presentation: Regulatory resource depletion impairs impression management and effortful self-presentation depletes regulatory resources. *Journal of Personality and Social Psychology*, *88*(4), 632–657.
- Wade-Benzoni, K. A., Tenbrunsel, A. E., & Bazerman, M. H. (1996). Egocentric interpretations of fairness in asymmetric, environmental social dilemmas: Explaining harvesting behavior and the role of communication. *Organizational Behavior and Human Decision Processes*, *67*(2), 111–126.
- Weick, K. E. (1995). *Sensemaking in organizations*. London: SAGE Publications.
- Welsh, D. T., & Ordóñez, L. D. (2014). Conscience without cognition: The effects of subconscious priming on ethical behavior. *Academy of Management Journal*, *57*(3), 723–742.
- Zhang, S., Cornwell, J. F. M., & Higgins, E. T. (2014). Repeating the past: Prevention focus motivates repetition, even for unethical decisions. *Psychological Science*, *25*(1), 179–187.
- Zhang, T., Gino, F., & Margolis, J. (2014). Does 'Could' lead to good? Toward a theory of moral insight Harvard Business School Working Paper No 14-118.
- Zhang, T., Gino, F., & Bazerman, M. H. (2014). Morality rebooted: Exploring simple fixes to our moral bugs. *Research in Organizational Behavior*, *34*, 63–79.
- Zhong, C. B., & House, J. (2012). Hawthorne revisited: Organizational implications of the physical work environment. *Research in Organizational Behavior*, *32*, 3–22.
- Zhong, C. B., Ku, G., Lount, R. B., & Murnighan, J. K. (2010). Compensatory ethics. *Journal of Business Ethics*, *92*(3), 323–339.
- Zhong, C. B., Liljenquist, K. A., & Cain, D. M. (2009). Moral self-regulation. In D. DeCremer (Ed.), *Psychological perspectives on ethical behavior and decision making* (pp. 75–89). Charlotte, NC: IAP.