



Morality Rebooted: Exploring Simple Fixes to Our Moral Bugs

**Ting Zhang
Francesca Gino
Max H. Bazerman**

Working Paper

14-105

April 21, 2014

Copyright © 2014 by Ting Zhang, Francesca Gino, and Max H. Bazerman

Working papers are in draft form. This working paper is distributed for purposes of comment and discussion only. It may not be reproduced without permission of the copyright holder. Copies of working papers are available from the author.

Running Head: MORALITY REBOOTED

**Morality Rebooted:
Exploring Simple Fixes to Our Moral Bugs**

ABSTRACT

Ethics research developed partly in response to calls from organizations to understand and solve unethical behavior. Departing from prior work that has mainly focused on examining the antecedents and consequences of dishonesty, we examine two approaches to mitigating unethical behavior: (1) *values-oriented* approaches that broadly appeal to individuals' preferences to be more moral, and (2) *structure-oriented* approaches that redesign specific incentives, tasks, and decisions to reduce temptations to cheat in the environment. This paper explores how these approaches can change behavior. We argue that integrating both approaches while avoiding incompatible strategies can reduce the risk of adverse effects that arise from taking a single approach.

Keywords: corruption; dishonesty; unethical behavior; interventions; structure; values

Following the corporate scandals at Enron, Halliburton, Worldcom, and several other notable firms, many called on business schools to reorient the moral compass of their students and on government to rethink policies. In response to these calls throughout the last decade, business schools have devoted more classroom time to ethics and policy makers have introduced new regulations. Despite these efforts, more scandals, such as the Madoff, Olympus, and the Libor cases, have surfaced year after year. These events may partially explain the growing interest in ethics research; consequently, the field of behavioral ethics has thrived (Bazerman & Moore, 2012; Messick & Bazerman, 1996; Treviño, Weaver, & Reynolds, 2006). Although we know far more now than we did before about the conditions under which individuals are likely to behave unethically (see Bazerman & Gino, 2012 and Monin & Jordan, 2009 for recent reviews), our current understandings are still primarily descriptive. We have not yet identified the main strategies groups and organizations can adopt to implement change and tested their effectiveness empirically (Kluver, Frazier, & Haidt, 2014; Moore & Gino, 2013). This paper responds to the challenge of advancing our current knowledge of unethical behavior from largely descriptive research to a framework aimed to reduce or even eliminate unethical behavior in organizations.

One of the robust findings of behavioral ethics research is that dishonesty is difficult to change due to three main reasons. First, individuals often engage in unethical behavior without the awareness that they are doing so (Chugh, Bazerman, & Banaji, 2005). Second, even when people recognize they are acting unethically, they fail to realize that social and situational forces are pushing them to cross ethical boundaries (Moore & Gino, 2013). Thus, rather than a stable trait, morality is both dynamic and malleable (Monin & Jordan, 2009): even if we care about being moral, most of us – under certain social or situational pressures – act unethically. Finally, unethical behaviors are often difficult to detect, especially when observers of the behavior

operate under motivated biases (e.g., people fear being harmed if they detect others cheating; see Gino, Moore, & Bazerman, 2009). This paper responds to the difficulty of reducing unethical behavior with a framework that takes into account the behavioral realities that challenge all of us.

Our goal is twofold. First, we identify approaches to mitigating unethical behavior based on empirical evidence from existing research in moral psychology and behavioral ethics. Second, we develop a framework for evaluating different strategies with prescriptive recommendations on how to reduce unethical behaviors. Drawing on the wealth of research on unethical behavior accumulated over the last decade that explains why individuals act immorally and the conditions that foster dishonesty, we dichotomize ethical fixes into two broad categories: *values-oriented* and *structure-oriented* approaches. Values-oriented approaches shift people's preferences to be moral, whereas structure-oriented approaches seek to design incentives, decisions, and tasks such that the unethical option is less tempting. Based on theory and empirical findings, we propose that adopting both values-oriented and structure-oriented approaches mitigates the risk of adverse effects from one strategy taken from a single approach. We discuss areas for future research and implications for theory, as well as business practice and policy.

WHY DO PEOPLE ACT UNETHICALLY?

In this paper, we use Jones' (1991) definition of unethical behaviors as those actions that have harmful effects on others people and are "either illegal or morally unacceptable to the larger community" (p. 367), comprised of groups, organizations or societies more broadly. Based on this definition, examples of unethical behaviors include—among others—violations of ethical norms or standards (whether legal or not), stealing, lying and cheating (Reynolds & Ceranic, 2007; Treviño et al., 2006). We use the term "unethical" to include cheating and acting dishonestly, immorally, and deceptively.

Traditional models in economics on crime suggest that individuals commit wrongful acts when the benefits of wrongdoing outweigh the costs (Becker, 1978; Jensen & Meckling, 1976; Prendergast, 1999). The distinguishing feature of these models is that individuals act out of self-interest and that they consciously choose to act either ethically or unethically, depending on the ratio of benefits to costs.

Although traditional versions of these “rational” models provide a parsimonious framework for understanding individual’s unethical actions, they do not focus on social attributes that a decision-maker might value, particularly the degree to which individuals value being honest. For example, whereas standard economic models would expect individuals to cheat to the maximum possible extent if there were no external costs, laboratory studies repeatedly show that most individuals cheat only a little bit—far from the maximum amount (Gino, Ayal, & Ariely, 2009; Mazar, Amir, & Ariely, 2008). And they cheat to the extent they can justify their actions to themselves, allowing them to maintain their self-image as a good person (Gino & Ariely, 2012; Mazar et al., 2008). Beyond considering the external costs and benefits based on anticipated punishments and rewards of acting unethically, individuals’ decisions to behave dishonestly also depend on the psychological costs and benefits of such behaviors (Messick & Bazerman, 1996). By incorporating psychological and behavioral approaches to the study of ethics, we highlight two notable departures from this traditional cost-benefit model that serve as foundations for our proposed strategies: 1) individuals have unstable preferences and utility functions, and 2) individuals exhibit bounded ethicality.

Unstable preferences for morality

Individuals have unstable preferences to act ethically in that the degree to which benefits of unethical behaviors seem enticing and costs unappealing depends on one’s temporary moral

self-concept and the normatively irrelevant characteristics of the surrounding environment (Greenwald, 1980; Griffin & Ross, 1991; Sanitioso, Kunda, & Fong, 1990). Nisan's moral balance model (1991) proposes that individuals compute a personal moral balance based on their behaviors that are morally relevant within a given time frame. Good deeds raise the balance, and bad deeds lower it. For instance, the anticipation of making decisions perceived as morally wrong (e.g., excluding minorities) motivates individuals to repair their moral image (e.g., expressing favoritism toward minority job candidates; Merritt et al., 2012). Furthermore, increasing people's moral self-regard can liberate them to act less ethically in the future (e.g., Effron & Monin, 2010; Mazar & Zhong, 2010; Monin & Miller, 2001). The implication is that while resisting the temptation to cheat may lead people to feel internally rewarded and to perceive themselves as having a good self-concept, the same act can license the individual to cheat more in the future. These findings partially explain how the same individual may inconsistently choose to act ethically in one situation, but unethically in another.

Additionally, individuals' moral self-concept and desire to be ethical depends not only on prior actions but also on temporal focus. Because individuals' values are more likely to influence predictions about the distant rather than the near future (Trope & Liberman, 2003), thinking about the future is more likely to elicit "should" decisions based on ethical intentions, whereas thinking about the present or near future is more likely to trigger "want" decisions that are more focused on immediate gain and self-interested wins (Tenbrunsel, Diekmann, Wade-Benzoni, & Bazerman, 2010). That is, when considering the tradeoff between long-term benefits of being an ethical person and short-term gains of acting dishonestly, the "want self" wins in the present, leading people to act unethically. However, when thinking about the distant future, individuals typically consider what they "should" do and are often overly optimistic that they can resist

temptations to cheat in the future. This temporal differentiation in activating the “want” self in the present and the “should” self in the future also explains why individuals with ethical values succumb to ethical lapses when confronted with temptations to cheat in the present.

Situational factors such as the social environment and individuals’ interpersonal relations also determine the extent to which ethics is factored into their decision (Moore & Gino, 2013). Based on assumptions in social identity theory, individuals generally feel rewarded when they identify with social groups and adopt their norms (Tajfel & Turner, 1986). The costs of breaking social norms of the groups and defying situational pressures can tempt even honest individuals to cheat or make uncharacteristically harmful decisions (Milgram, 1974). Furthermore, individuals are more likely to cheat when other in-group members are acting unethically, but less likely to cheat when out-group members are being dishonest (Gino et al., 2009). Thus, the desire to be moral is highly dynamic and malleable, rather than stable in that it reacts and molds to the situational context (Monin & Jordan, 2009). Because morality is context dependent and it is socially constructed (Moore & Gino, 2013), dishonesty does not only pertain to those who do not value ethics, but also those who do value morality but find themselves in situations with social pressure to act unethically.

Bounded Ethicality

While many unethical acts are intentional in that actors are aware they are violating ethical standards when engaging in the act (e.g., stealing, cheating, and lying), many actions are unintentional. Individuals would likely change or retract these decisions upon realizing their ethical implications. Given the limits of the conscious mind and the pervasiveness of the unconscious one (Wegner, 2003), robust empirical evidence has highlighted how unethical behavior often results from unconscious rather than conscious actions (Banaji, Bazerman, &

Chugh, 2003; Chugh et al., 2005). This research has suggested that individuals' ethicality is bounded (Banaji et al., 2003). Decisions made under bounded ethicality are often outside of the actor's own awareness and inconsistent with the actor's consciously held ethical values (Bazerman & Moore, 2012). These unintentional acts of wrongdoing made without moral awareness fall outside of the traditional cost-benefit framework since individuals do not even recognize the ethical issues under consideration.

People carry both positive and negative attitudes that are outside of conscious awareness, and these implicit attitudes can lead to unintended discriminatory behaviors (Banaji & Greenwald, 2013; Chugh et al., 2005; Greenwald & Banaji, 1995). For instance, rather than arising from intentional exclusion of social out-groups, some discriminatory decisions arise from implicit favoritism of in-group members (Messick & Bazerman, 1996). Bounded ethicality also leads to other types of ethical failures, including individuals' tendency to over-claim credit for joint work without realizing that they are doing so (Epley, Caruso, & Bazerman, 2006; Ross & Sicoly, 1979), and over-discount the future in ways that harm the environment (Tenbrunsel et al., 2010) and future generations (Wade-Benzoni, 1999).

Because implicit biases are particularly persistent and difficult to change, most research on ethics with implications on improving morality has focused on situations in which individuals are morally aware. And even in that domain, little research has studied potential fixes to ethical failures. In this paper, we draw on existing research to propose and evaluate fixes aimed to mitigate both explicit and implicit forms of dishonesty.

APPROACHES TO FIXING OUR MORAL BUGS

Based on research in moral psychology, philosophy, neuroscience, economics, and behavioral ethics, we identify two main approaches to target unethical behavior: *values-oriented*

and *structure-oriented* approaches. *Values-oriented* approaches shift individuals' preferences to be moral by reminding them of their moral values or the moral values of the group or organization when making decisions with ethical implications. In contrast, *structure-oriented* approaches, seek to reshape the structure of the incentive, decision, or task in order to reduce the degree to which acting dishonestly is tempting, or in some cases, even an option. As an illustration of these two approaches, consider the following example. To prevent students from peeking at other students' answer sheets during tests, the values-oriented method shifts students' preferences to be more ethical by reminding students of their ethical identity or of the ethical standards in the classroom (e.g., signing an ethics code, having students create their own code of conduct, or putting up the mission statement of the school on the classroom walls). In contrast, structure-oriented solutions entail changing the design of the exam so that students are less tempted to cheat (e.g., randomizing order of exam questions and choices, forcing students to sit further apart, or simply increasing the penalty of being caught).

In the following sections, we explore how each of these approaches impact behavior and suggest that organizations make themselves more vulnerable to adverse effects when they consider the two approaches in isolation rather than in concert. Given that these approaches impact behavior through non-overlapping mechanisms and have different sets of limitations, we propose that individuals and organizations that can consider both approaches are better positioned to mitigate the adverse effects of taking any one given approach.

Values-oriented approach: Shift preferences towards morality

Given the malleable and dynamic nature of morality, one approach to curbing dishonesty entails shifting people's preferences to be moral. That is, instead of changing the external environment to shape behavior by altering the structure of incentives, tasks, or decisions that

individuals face, a values-oriented approach directly targets the individuals' internal desire to be ethical. By reminding individuals of their ethical self-concept or the ethical norms of their organization, incongruences between values and unethical behaviors increase the psychological costs of wrongdoing. These values-oriented strategies have been shown to target a wide variety of unethical behaviors, including—but not limited to—cheating, stealing, and lying (Gino & Mogilner, 2012; Mazar et al., 2008; Shu, Mazar, Gino, Ariely, & Bazerman, 2012).

One long-term solution of reducing dishonesty at the individual level is reminding individuals of their personal moral self-concept (i.e., how individuals view and perceive themselves) since it is a crucial determinant of future engagement in unethical behavior (Aquino & Reed, 2002). Having a strong moral self-concept or identity compels individuals to behave ethically (Colby & Damon, 1994; Oliner & Oliner, 1988). Aquino and Reed (2002) distinguish between two dimensions of moral identity: internalization (i.e., the degree to which various moral traits are central to the self-concept) and symbolization (i.e., the degree to which individuals' actions in the world express these traits). These dimensions have been found to predict several moral behaviors, including self-reported volunteering and individuals' willingness to minimize harm (Aquino & Reed, 2002; Reed & Aquino, 2003). Thus, creating a connection between past moral behaviors and individuals' moral self-concepts—either by reminding individuals of past costly behaviors that signal to individuals their moral selves or by reflecting on previous moral deeds abstractly—helps individuals engage in ethical behaviors consistent with their past moral actions (Conway & Peetz, 2012; Gneezy, Imas, Brown, Nelson, & Norton, 2012).

In addition to targeting individuals' moral self-concept and motivations, another method of reducing unethical behavior is to raise the salience of ethical values espoused by the group.

Because individuals generally feel rewarded when they identify with social groups and adopt their norms and values (Tajfel & Turner, 1986), organizations can evaluate both *descriptive* (e.g., actions that leaders take) and *injunctive* norms (e.g., explicit messages on codes of conduct) within the organization.

Although some values-oriented strategies consciously bring moral values into the foreground, values-oriented strategies need not consciously direct individuals' attention to their moral values. Another method of increasing ethical behavior is to prime individuals' desires to be ethical with concepts related to morality and avoid those that prime unethical motivations. For example, given that individuals generally perceive a strong relation between children and concepts such as purity and innocence (Wright, Hanoteau, Parkinson, & Tatham, 2010), mere exposure to childhood-related cues activates moral goals and leads individuals to behave less unethically and more prosocially (Desai & Gino, 2012; Gino & Desai, 2012). Furthermore, shifting the individuals' focus away from thoughts that can trigger immoral behaviors helps individuals avoid unethical acts. For example, shifting thoughts away from money to thoughts about time also reduces cheating (Gino & Mogilner, 2012) because it raises self-reflection. Based on these findings, organizations could impact a wide range of unethical behaviors by thinking more carefully about how concepts of time (in terms of long-term focus), self-reflection and innocence can be primed through language used in company policies or documents (e.g., in communications from leaders in organizations).

Limitations. Values-oriented approaches on their own, however, may be ineffective at counteracting particularly strong forces that tempt individuals to cheat: specific sources of unethical behavior that organizations are likely eager to address. For example, organizations commonly turn to values-oriented strategies such as creating mission statements and ethics codes

to curtail dishonesty; however, evidence from field studies within organizations provide mixed findings about the effectiveness of these messages (Cleek & Leonard, 1998; McCabe, Trevino, & Butterfield, 1996; Stevens, 2008). Historic cases of this values-based approach also provide evidence of mixed success. Although both Johnson & Johnson and Enron presented codes of conduct to their employees, the companies took two divergent paths: Johnson & Johnson created an ethical culture that enabled it to make the difficult decision to voluntarily recall its own product during a cyanide-tainting crisis that cost the company \$100 million, whereas Enron created a culture of greed that led to one of the biggest accounting scandals and corporate bankruptcies to date (Bazerman & Tenbrunsel, 2011). The conclusion from these examples is not that values-oriented strategies are unreliable or do not work, but that the degree to which these interventions can specifically target a wide range of unethical behaviors depends on a number of structural elements in place, such as how the ethics codes are used, whether they are made salient at the time of decision, the extent to which organizational leaders reinforce them with their behavior, and how the overall message interacts with other competing goals and messages of the organization. Without these structural considerations, values-oriented strategies alone may not be able to change the pervasiveness of particularly resilient forms of unethicality.

We note that this set of fixes generally targets intentional acts of unethical behavior in which individuals are morally aware and understand the external costs and benefits of cheating. Thus, values-oriented strategies are likely ineffective at addressing cases of bounded ethicality. That is, because a values-oriented approach impacts behavior by making inconsistencies between values and actions more salient, this approach relies on the assumption that individuals recognize when their actions violate these ethical standards. Reminders of moral values are ineffective when targeting unethical behaviors in which perpetrators are unwittingly breaking these

standards. As we describe in the following sections, a combination of values-oriented and structure-oriented strategies is needed to bring these individuals into a state of ethical awareness and help them avoid past breaches of ethics.

Risk of adverse effects. When structural considerations are not taken into account, organizations run the risk that these values-oriented messages contradict moral messages implied through existing structural features and actions. For instance, organizations that promote ethical mission statements while failing to adjust unrealistic goals that routinely place employees in ethical dilemmas. These incompatibilities may lead to unintended consequences.

The exact nature of these consequences depends on whether employees can justify their unethical actions as moral. If employees can generate justifications for these unethical actions, they may become morally disengaged by convincing themselves that these unethical actions are in fact morally permissible (Bandura, 1999; Bandura, Barbaranelli, Caprara, & Pastorelli, 1996; Shu, Gino, & Bazerman, 2011). If employees cannot justify their immoral actions but are still compelled to act unethically (e.g., due to stringent requirements organizations set to meet unrealistically high goals), individuals may convince themselves that these values-oriented messages can be ignored. By failing to examine structural elements that are incompatible with their values-oriented messages, organizations may permit negligence of values, worsening the extent to which employees engage in cheating and even unintentionally promoting unethical behavior in areas where employees had not acted unethically. Furthermore, even if these moral messages can persuade individuals to resist the temptation to cheat, continually exercising self-restraint can deplete self-regulatory resources and promote cheating behavior when individuals are confront with new temptations (Gino, Schweitzer, Mead, & Ariely, 2011).

Additionally, simply considering the content of moral messages without considering the structural components of *when* and *how* these moral messages are relayed may inadvertently engrain unethical behaviors more deeply. For example, consider the numerous times in which we fill out forms to provide important information that others rely on to be true (e.g., tax forms, job applications, and credit card applications). Only after providing important information do we sign our names at the bottom to confirm that we have been truthful. But why do we sign at the end? Signing ethics codes before facing the opportunity to cheat—rather than after—raises the salience of morality at the appropriate moment, and as a result, reduces dishonesty (Shu et al., 2012). In their multi-study paper, Shu and colleagues (2012) conducted a field experiment at an automobile company and found that those who signed at the bottom of a mileage report form were less honest—they reported driving approximately ten percent fewer miles on average compared to those who signed an ethics code at the top of the insurance form. Despite this evidence, form creators still use traditional document in which people act in unethical ways and then are asked to sign to verify the truth of its contents. In fact, form makers might even unwittingly encourage individuals to escalate their unethical actions by having individuals sign at the bottom after unethical commitments have been made. Thus, failure to consider structural approaches with values-oriented strategies can even exacerbate unethical decisions.

Structure-oriented approach: Reduce specific acts of unethical behavior

An alternative approach toward improving ethics is not relying on individuals' internal desire to act ethically, but instead structuring the incentives, task, or set of choices to reduce or even eliminate the temptation to act unethically (Becker, 1978; Messick & Bazerman, 1996; Thaler & Sunstein, 2008). Unlike the values-oriented approaches that generate psychological cost of values-inconsistent actions, structure-oriented approaches impose external costs of acting

unethically, make dishonesty more inconvenient to implement, or altogether remove unethical options from the set of possible actions. Thus, structure-oriented strategies tend to target particular forms of unethical behavior by decreasing the temptation to cheat in the situation itself. For example, raising the penalty of tax evasion or setting defaults on taxes owed are both structure-oriented interventions that specifically target dishonest reporting of taxes.

Based on traditional economic models, research on incentives has focused on increasing the size of punishment or probability of being caught to deter unethical behavior (Becker, 1978). Looking beyond incentives, Sunstein and Thaler (2008) proposed a different set of structure-oriented strategies to improve decision making: choice architecture. That is, instead of changing financial incentives, restructuring choice sets and the accessibility of these choices aims to nudge people to make wiser decisions while simultaneously preserving individuals' freedom of choice (Thaler & Sunstein, 2003; Thaler & Sunstein, 2008). Much of the research on choice architecture has been aimed to improve self-control problems and eliminate decision biases such as optimizing saving decisions, increasing organ donation rates, or promoting healthier habits. However, we believe that many of the same principles apply to reducing unethical behavior, particularly as individuals who have more difficulty regulating their behaviors effectively also tend to engage in more harmful behaviors (Cohen, Panter, Turan, Morse, & Kim, in press). Because choice architecture often nudges individuals to adopt ethical choices or avoid unethical ones outside of individuals' own awareness, these changes may reduce the extent to which behaving ethically is cognitively demanding, thus freeing resources for other tasks (Thaler & Sunstein, 2003; Thaler & Sunstein, 2008).

A classic example of choice architecture is employing default options as subtle nudges to help individuals make more desirable choices. Through a well-known study conducted in the

context of organ donations, Johnson and Goldstein (2003) show that default choices can have tremendous impact on whether people consent to donate their organs. In countries where citizens must “opt in” as an organ donor, donation rates were no greater than 30 percent; however, in countries where citizens must “opt out” of donating their organs, nearly every citizen consented to donate. Such defaults could also have large implications within the ethics domain. For example, given the amount of tax evasion every year, U.S. policy makers could simplify the reporting process on taxes by adapting the tax payment process in European countries. For example, Danish tax payers receive a preliminary tax statement with their estimated income for the prior year listed on the form and are only required to take action if the estimated amount is incorrect (SKAT, 2005). Implications of defaults research include not only setting ethical defaults for decisions without defaults, but also re-examining existing defaults. Organizations have implicitly created defaults that affect the behavior of employees. Auditing these defaults is an activity worth the attention of an organization striving for more ethical conduct.

Although structural interventions entail redesigning tasks to mitigate the temptation to cheat, they need not dramatically alter the design of the task. Even subtle shifts in the framing of tasks (e.g., depicting goals in terms of gains instead of losses) can reduce unethical behaviors. In a series of experiments, Kern and Chugh (2009) manipulated the framing of a decision in terms of losses (there is a 75% chance of losing an acquisition) or gains (there is only a 25% chance of gaining the acquisition). Because individuals are averse to losses, those in the loss-frame condition were significantly more likely to engage in unethical behavior, such as gathering “insider information” and lying to others, than were participants in the gain-frame condition (Kern & Chugh, 2009). These findings show that whereas some structural interventions entail removing the opportunity to cheat, or making cheating more inconvenient, other structural

interventions can indirectly shift the temptation to cheat without changing the underlying nature of the task.

Limitations. Because structure-oriented interventions are targeted toward a specific form of unethical behavior, they rarely alter behavior outside of the area of specified treatment. Just as using defaults for organ donation likely does not influence individuals to engage in more prosocial behavior in other domains (i.e., donating blood or volunteering at a non-profit), using defaults to encourage honest reporting on taxes likely does not discourage other types of unethical behaviors (i.e., inflating travel expenses to obtain larger reimbursements).

Furthermore, although there is growing evidence that structure-oriented interventions can drastically influence behavior (Johnson & Goldstein, 2003; Thaler & Sunstein, 2008), this approach may have limited impact if values of the individual are not aligned with the values of the structural intervention. For instance, although countries with presumed consent systems for organ donation have greater proportion of individuals who are organ donors, they do not necessarily have higher procurement rates of organs. In both informed and presumed consent countries, doctors discussed the donation decision with family members and did not authorize the donation if family members objected, leaving some presumed countries to have similar—or even lower—organ procurement rates as compared to informed consent countries (Boyarsky et al., 2012). In the context of savings decisions, default options to save do not always lead to increased savings rates. For instance, setting U.S. Savings Bonds as the default option for tax refunds did not increase allocation towards savings as these low-income refund recipients had already made plans to use their money. Together, these findings suggest that simply relying on structural changes to shift behavior without promoting the values behind the changes may not achieve

desired results, especially if the values of the individual do not align with the implicit values of the structural change.

Risk of adverse effects. Neglect of values-oriented approaches may also generate perverse consequences if the structural intervention precludes individuals from considering the issue as an ethics-related decision. For example, people who otherwise would honestly report their income on taxes may be more likely to underreport their income if the default provided is mistakenly lower than their actual income. Additionally, Tenbrunsel and Messick (1999) found that sanctioning systems with a small probability of being caught and small punishments were less effective than having no sanctioning systems at all in reducing dishonest behaviors. In their experiment, decision makers under these weak sanctioning systems tended to frame their decisions as business-related, whereas those under no sanctioning systems at all were actually more likely to frame the decisions as ethics-related and thus make more ethical decisions (Tenbrunsel & Messick, 1999). This research suggests that using incentives to highlight the negative side to unethical behavior could lead to even more wrongdoing as doing so may prevent individuals from perceiving their decisions as ethically-relevant.

Additionally, without shifting individuals' preferences to be moral, relying on structure-oriented strategies may simply lead individuals to engage in alternative forms of unethical behavior. For example, randomizing the order of test questions may prevent students from copying each other's test answers during the exam, but may not prevent students from alternate forms of cheating such as discussing answers during breaks or hiring stand-ins to take exams on the behalf of students. Because it is not possible to control all aspects of the environment to remove temptations to cheat, over-reliance on structure-oriented changes may also leave individuals ill-prepared to resist temptations to cheat when they do encounter temptations to act

unethically. Research on general self-control outside of the moral domain has shown that individuals who repeatedly exercised self-control over time were more adept at self-regulation than those who did not (Muraven, Baumeister, & Tice, 1999). Furthermore, individuals who built a strong ethical identity were more immune to temptations to cheat (Gino et al., 2011). Thus, over-reliance on purely structure-oriented changes may also inadvertently de-emphasize the importance of ethical awareness, prevent individuals from learning to exercise self-restraint over time, and degrade individual's ability to resist temptations on their own without the help of structure-oriented changes.

A DUALISTIC APPROACH

Given the strengths and weaknesses of values- and structure-oriented approaches on their own, we propose that incorporating both approaches can compensate for each approach's unique set of limitations and dampen the risk of adverse effects. Individuals, organizations, and policy makers can adopt this approach in two ways: 1) audit for incompatible strategies that increase the risk of adverse effects, and 2) create integrative strategies that redesign the structure of the decision or task to remind individuals of their ethical values. In the following section, we provide concrete examples of solutions that adopt each of these approaches based on recent research to demonstrate this point. Additionally, we note that using a dualistic approach can be particularly useful for identifying and addressing issues of bounded ethicality in which individuals are not aware that they are violating their ethical standards—a problem that has been traditionally difficult to pinpoint and change given the actors' lack of awareness.

Remove incompatibilities

When implementing either structure- or values-oriented interventions, failure to audit for incompatible interventions can lead to unintended consequences. For example, simultaneously

emphasizing the ethicality of employees while also installing video cameras to monitor theft sends incompatible messages. Discussion of employees as ethical signals the organization's trust in its employees' values, but the installation of monitoring equipment directly undermines this message, creating the risk of employee disengagement with or even reactance to the moral message.

Just as it is important to identify incompatibilities while implementing multiple structure- and values-oriented interventions, it is equally important to identify how introduction of new interventions may conflict with existing interventions. For example, auditing for other goals that counter value-based messages is critical to ensure values-based messages are not overlooked or devalued. Because individuals are more likely to resort to unethical means when goals are too difficult (Schweitzer, Ordóñez, & Douma, 2004; Welsh & Ordóñez, 2014), creating an ethics code may be less effective if managers have simultaneously set unrealistically high goals that may pressure employees to engage in more unethical behavior. These findings explain why setting a high sales goal of \$147 an hour led Sears, Roebuck car mechanics to overcharge customers and repair unnecessary fixes even when the company had intentions to promote ethical behavior (Bazerman & Tenbrunsel, 2011).

Furthermore, an organizations' approach to ethics as either aspirations to be advanced (promotion-focus frame) or as violations to be prevented (prevention-focus frame) can interact with the regulatory focus that other organizational goals induce within individuals: whether their task focuses on reaching a desired goal (promotion-focus) or avoiding an undesired goal (prevention-focus) (Gino & Margolis, 2011). In a series of laboratory studies, promotion-focus individuals who also read a promotion-focus ethics code were the most dishonest, whereas prevention-focus individuals who read a prevention-focus ethics code were the least dishonest.

These findings provide a cautionary tale that an organization's promotion-focused ethical goals as described by its mission statement, ethics code, policies, or incentive schemes can encourage unethical behavior if other goals in the organization are also promotion-focused (Gino & Margolis, 2011). Thus, removing existing structure-oriented strategies that compete against values-oriented messages can more successfully mitigate unethical behaviors.

Incorporate integrative strategies

In addition to eliminating incompatible structures and values, integrative strategies aim to combine the strengths of both strategies by incorporating individuals' values in the structure of the decision or task. Unlike structural changes that generally do not raise awareness of individuals' values, these changes are specifically designed to encourage consideration of values at the moment of temptation.

Align values and timing of decision. As noted in the discussion of values-oriented approaches, simply making moral standards explicit is often insufficient. Considering a structure-oriented approach in designing values-oriented interventions can improve ethics. For example, organizations too often write mission statements and ethics codes without considering how these messages fit with the decisions and tasks individuals in the organization face. Adjusting specifically *when* these standards are made salient can make a significant difference on the prevalence of dishonesty.

In particular, we propose that values-based messages are more effective if they are implemented before, rather than after, the moment individuals make decisions carrying important ethical considerations. Just as strategically placing salads at the beginning of the cafeteria line increases healthy eating (Just & Wansink, 2009), considering ethical implications of a decision prior to other aspects of the decision may reduce the propensity of individuals to overlook these

ethical considerations. In one simple intervention, signing ethics codes before completing tax and insurance forms—as opposed to after—reduced unethical behavior (Shu et al., 2012). These results suggest that emphasizing ethical considerations first can also anchor individuals to consider the ethical implications of subsequent decisions.

Furthermore, increasing the time between the decision and the enactment of that decision may reduce unethical behavior. Research shows that the degree to which individuals are future- or present-oriented shapes how individuals construe decisions (Liberman, Sagristano, & Trope, 2002; Liberman & Trope, 1998; Trope & Liberman, 2003), including decisions made in the moral domain (Conway & Peetz, 2012). For instance, people express greater disapproval of unethical behavior when they are making connections between how their current selves relate to their future selves than when they do not see the continuity between their current and future selves (Hershfield, Cohen, & Thompson, 2012). When considering future states, individuals think in a more high-level, abstract way such that they are more likely to make “should” decisions; however, when thinking about the present, they are more detail-oriented in that they think more about the visceral experience of the choices and as a result are less able to resist “want” choices (Rogers & Bazerman, 2008).

Applied to the domain of ethics, individuals are less likely to consider unethical options when those decisions are enacted in the future than when enacted in the present (Tenbrunsel et al., 2010). Structure-oriented strategies take advantage of these value inconsistencies across time (Ainslie, 1992; Laibson, 1997; Loewenstein & Prelec, 1992; Milkman, Rogers, & Bazerman, 2008; O'Donoghue & Rabin, 1999; Phelps & Pollak, 1968). For instance, phone calls encouraging citizens to vote in the future by facilitating the formation of a voting plan in the present increased voting participation by 9 percent (Nickerson & Rogers, 2010). Similarly,

employees who received a specific prompt to write down both a date and a time they intend to receive their flu vaccinations had a 4.2 percentage point higher vaccination rate than those who simply wrote down a date or received a reminder (Milkman, Beshears, Choi, Laibson, & Madrian, 2011). The analog of these interventions in the moral domain is that individuals can act more ethically by making detailed plans in the present to confront or avoid anticipated temptations to act dishonestly in the future (e.g., people can commit to buy a music album on its release date to avoid the temptation to download a music album illegally in the future). Importantly, Milkman et al.'s study (2011) suggests that even private commitments are effective, making this strategy particularly applicable to ethical goals that are more sensitive and personal compared to other goals.

Relatedly, commitment contracts, or incentivized versions of these self-promises that financially “lock in” decisions made for the future, also integrate values- and structure-oriented approaches to enact change (Beshears, Choi, Laibson, Madrian, & Sakong, 2012; Thaler & Benartzi, 2004). Because these contracts often impose an external fine if the commitment is not met at the future date, individuals feel more obligated to stick to the “should” decisions to which they had previously agreed, even if the “want” decision becomes more attractive as the time of the decision draws nearer (Ariely & Wertenbroch, 2002; Ashraf, Karlan, & Yin, 2006). Although prior research on commitment contracts has focused on mechanisms to help individuals commit to “should” decisions outside of the moral domain (i.e., saving decisions) (Thaler & Benartzi, 2004), we believe that this research is also applicable to ethical behaviors. For example, given that time-pressure influences unethical behavior (Shalvi, Eldar, & Bereby-Meyer, 2012), individuals can indirectly curb dishonesty by committing to start tasks earlier (e.g., commit to study for an exam well before the exam date and agreeing to pay a cost if the commitment is not

met in order to discourage cheating). While these commitment contracts may not target unethical behavior directly, they can target key drivers of unethical behavior.

Integrating values- and structure-oriented approaches is not only applicable to interventions at the individual-level, but also at the societal level through policy design. Since individuals are more likely to make “want” decisions that may be less ethical in the present and “should” decisions that may be more ethical in the future, the principles underlying the concept of future lock-ins can allow policy makers to promote and implement “should” policies that are unpopular in the present. Rogers and Bazerman (2008) demonstrate that when considering a policy to increase gas prices that would be voted on as soon as possible and then implemented in two years, individuals were more likely support the policy when the distant future was emphasized (implemented in two years) than when the near future was emphasized (voted on as soon as possible). Simply framing policies as decisions for the future can encourage people to take on immediate, short-term costs for long-term gains—an issue that is particularly relevant in the moral domain.

These findings demonstrate that integrating remarkably simple structural considerations into values-oriented changes can have large implications on the ethicality of responses. By simply taking into account the timing of when individuals view values-oriented messages and commit to decisions, these interventions can more specifically target intended unethical behaviors.

Joint versus separate evaluation. Another possible fix that integrates both values- and structure-oriented approaches involves evaluating decisions jointly rather than separately. Unlike some structure-oriented fixes that intended to eliminate the need to consider values, this strategy uses structure to bring values to the foreground that otherwise would have been ignored. Given

the difficulty in eliminating individuals' implicit biases as individuals are by definition unaware that these biases exist, joint evaluation could be a powerful solution to mitigating unintentional unethical behaviors. In the moral domain, this fix has been examined as a solution to discriminatory behaviors that commonly occur in organizations. For instance, a recent study explores possible interventions organizations can implement to overcome gender biases in hiring (Bohnet, Bazerman, & Van Geen, 2012). In this research, evaluators who judged candidates separately were more likely to neglect data on candidates' past performance and instead hire male candidates for math tasks and females for verbal tasks based on gender stereotypes, even though gender was not predictive of performance (Bohnet et al., 2012). However, those who evaluated a male and female employee did not exhibit these preferences consistent with gender stereotypes and instead made decisions based on the higher performing employee in each of the domains (Bohnet et al., 2012).

Additionally, joint evaluation has been found to help individuals make more ethical judgments. In a series of laboratory studies, Paharia and colleagues (Paharia, Kassam, Greene, & Bazerman, 2009) show that when making separate evaluations, people were more condemning of others that engaged in unethical behavior directly than those that engaged in unethical behavior indirectly; however, joint-evaluation reversed these judgments. These inconsistencies in decisions between joint and separate evaluations results from differences between System 1 thinking that relies on intuitions and System 2 thinking that is more deliberate and rational. When individuals evaluate separately, they do not have another point of comparison and often rely on their gut reaction to make judgments about the situation in front of them. However, when individuals are evaluating jointly, they are triggered to make comparisons that slow down their

judgments and that rely more on System 2 thinking. Thus, framing decisions jointly induces more deliberate thinking that can help overcome implicit biases.

Furthermore, joint evaluation can lead to more reason-based utilitarian decision making. When participants were either asked to evaluate a situation in which they could flip a switch to divert a speeding train from killing three individual to killing one, or push a man off a bridge to stop the train from killing five individuals, participants making separate evaluations were more likely to flip the switch to save three lives than push the man off the bridge to save one. However, participants were more likely to make the utilitarian decision when these situations were presented jointly to participants (Bazerman, Gino, Shu, & Tsay, 2013; Cushman, Tsay, Greene, & Bazerman, 2013).

Policy makers can also incorporate joint valuation by bundling policies with offsetting costs and net benefits in order to encourage support for policies with greater short-term costs for long-term gains (Milkman, Mazza, Shu, Tsay, & Bazerman, 2012). Because losses loom larger than gains psychologically (Tversky & Kahneman, 1991), policies that have higher immediate costs are more likely to be rejected when presented separately than when presented with another policy that partially offsets these costs (Milkman et al., 2012). For example, in a series of laboratory studies, participants were either asked about their support for one of two bills separately or for the policy bundle with both bills. When participants were asked to evaluate each separately, neither bill received more than 50 percent of the votes, meaning that neither bill would have been enacted into law; however, when presented as a bundle, 71 percent of participants voted for the bill, generating a potential net benefit for society compared to the status quo (Milkman et al., 2012). By incorporating values and structure, these changes, allow policy makers to target a wide range of behaviors.

Encourage mental and social contemplation. Another structure-oriented intervention infused with consideration of values is giving individuals more time to simulate events mentally and contemplate their actions. Such contemplation can also take the form of social exchanges in which individuals discuss their moral issues with others. Importantly, social norms heavily determine the impact of these conversations such that under moral norms, social contemplation decreases instances of lying, whereas under norms of self-interested behavior, social contemplation actually increase instances of lying (Gunia, Wang, Huang, Wang, & Murnighan, 2012). These findings suggest that open discussions of ethical issues are effective if the social norms of acting ethically are in place; otherwise, these social contemplations can be counterproductive, particularly when the organizational norms are based on self-interest.

When social contemplation is not possible, simply leaving more time for deliberation also generates greater honesty, especially when people have fewer justifications for dishonesty (Shalvi et al., 2012). In a laboratory experiment, participants who were incentivized based on the first die roll out of three opportunities to roll (i.e., 1=\$1, 2=\$2, etc.) lied more when working under a time constraint than when under no time pressure. Allowing more time is particularly beneficial for cases in which cheating cannot be justified—when participants could only roll the die once, those without a time constraint did not lie at all, whereas those under a time pressure cheated nearly just as much as they did in the previous case. These findings show that some individuals' initial reaction was to cheat, and only with more time did they restrain themselves from acting unethically. Furthermore, asking people to mentally simulate events by closing their eyes has been shown to help individuals judge immoral behaviors as more unethical and also decrease self-interested behavior (Caruso & Gino, 2010). Similarly, asking people to “think hard” about their decisions can also deter unethical behavior. Compared to individuals who were

asked to make decisions immediately, individuals were less likely to lie if they had the opportunity to spend time to mentally contemplate their future actions on their own or engage in social contemplation by conversing with others about the moral issue (Gunia et al., 2012).

Darley and Batson (1973) provide evidence of the perverse effects of time constraints. In their experiment, subjects in a hurry to reach their destination were more likely to pass by an individual slumped in the doorway without stopping to help. Dispositional factors such as an individuals' religiosity as well as other situational factors such as exposure to concepts related to purity (i.e., whether the person was on their way to give a short talk about the Good Samaritan fable or a topic unrelated to helping behaviors) did not moderate the primary effect (Darley & Batson, 1973). Thus, when people have more time, they are more likely to help others (Darley & Batson, 1973) and less likely to commit unethical acts (Kern & Chugh, 2009; Shalvi et al., 2012).

Connect self-concept to environment. Another approach that seeks to embed a values-oriented approach in designing a structure-oriented intervention includes implementing changes in the environment to induce self-awareness and highlight the link between behaviors and the moral self. For instance, individuals are less likely to cheat if they are exposed to a mirror while facing the temptation to cheat (Batson, Thompson, Seuferling, Whitney, & Strongman, 1999; Diener & Wallbom, 1976; Vallacher & Solodky, 1979). Similarly, decreasing the feeling of anonymity can also reduce unethicality. Being exposed to an image with a pair of eyes decreased littering in a university cafeteria (Ernest-Jones, Nettle, & Bateson, 2011) and increased contributions to an honesty box collecting money for drinks in a university coffee room (Bateson, Nettle, & Roberts, 2006). These findings show that establishing the link between identity and behavior need not be constrained to the messages and social norms that

organizations set—the design of physical spaces where employees work can also trigger the extent to which employees make the connection between their actions and self-concept.

GENERAL DISCUSSION

Based on the pervasive nature of unethical behavior, ethical goals seem to be sitting in the back seat despite the growth in knowledge about the antecedents and consequences of unethical behavior. This paper discusses the power of subtle changes and shows how they can point our moral compass toward a more ethical direction. Taking into account that our morality is malleable and that wrongdoing does not require awareness, researchers have taken two main approaches to rebooting morality: 1) a values-oriented approach that shifts individuals' preferences to be ethical by reminding them of their moral ideals or the moral ideals of the group or organization, and 2) a structure-oriented approach that redesigns the decision itself such that the temptation to cheat is no longer appealing during the moment of the decision.

While these two approaches both target unethical behaviors, they operate very differently. Because values-oriented fixes target individuals' desires to be more ethical and stresses the importance of ethical values more generally, these fixes bring ethicality into awareness for the individual. However, values-oriented interventions may be less effective if there exist countervailing structural forces that tempt individuals to act unethically (e.g., existing goals are framed as losses or difficult to reach). In contrast, because structural interventions target specific forms of dishonesty, they may have limited influence in other domains where the intervention is not present. For example, we posit that although joint evaluation discourages discriminatory decisions based on gender, this intervention would not impact cheating on taxes—a form of dishonesty that is outside the scope of impact for this structure-oriented intervention.

Given the limitations of each approach, we encourage organizational scholars to consider a dualistic approach that integrates strategies from each approach while avoiding incompatible ones. We highlight that a dualistic approach may even be helpful for situations in which individuals are ethically unaware. By identifying these approaches to reducing unethical behavior, we call for a shift in focus in behavioral ethics research. Over the last two decades, we have learned a lot about the antecedents and consequences of unethical behavior based on research across various disciplines, including moral psychology, organizational behavior, economics, and management more generally. Although we recognize the important insights this research has identified, we argue that it is now time to put more effort into the understanding how values-oriented and structure-oriented changes can interact to reduce unethical behavior in the workplace and in society more generally. By examining how to best introduce the fixes we discussed in this paper, investigating the necessary conditions for the fixes to produce desired effects, and exploring how organizations can resist the temptation to focus on a single approach when a dualistic approach may be more appropriate, we hope future research will solve the ethical challenges that organizations and their managers so commonly face.

Limitations and areas for future research

The prescriptions based on ethics research we discussed are grounded in economic and psychological models of unethical behavior. One limitation that applies to all of these interventions is that even employees who support the underlying values of these fixes may resist change if they perceive these interventions as forms of coercion. More research is needed to understand how change agents can minimize the extent to which these interventions are perceived as imposed change. Furthermore, if and when employees do resist, more research is needed to understand effective strategies that not only minimize the downside of resistance on

ethical decision making but also reconstruct these acts of resistance to benefit the organization in the long run (Ford, Ford, & D'Amelio, 2008).

When considering how to implement interventions targeted at reducing or eliminating unethical behavior, there are three critical questions that have remained unanswered. First, when should these interventions be introduced? Findings from research on the timing of ethics codes suggest that that some interventions should occur right before or during the moment of temptation (Shu et al., 2012), whereas other strategies depend on timing interventions well before the moment of temptation (Milkman et al., 2011; Nickerson & Rogers, 2010; Rogers & Bazerman, 2008; Thaler & Benartzi, 2004). More research is needed to understand how the effectiveness of interventions varies with the timing of intervention. Secondly, what effects do the same interventions have over time? Are one-time interventions enough, and if not, how does repeated exposure to the same intervention impact its effectiveness over time? For example, how do people respond to ethics codes with every additional ethics code they encounter? Lastly, what happens when these interventions are removed—will behavior resume back to pre-intervention levels or persist? These questions reveal that the structural considerations of *how* inventions are implemented are just as important as *what* interventions are implemented.

Relatedly, more future research is needed to understand how these strategies differ in *magnitude* or the extent of impact across these varying approaches and strategies. In particular, the magnitude of impact can be divided into three main components: the size of the treatment effect given a particular intervention (*strength*), the length of time these effects hold over the course of the intervention (*duration*), and the length of time these effects hold after the intervention (*persistence*) (Hunt & Rogers, 2013). Given the lack of comparability in results of research employing varying methodologies in different environments, we have little evidence to

compare the effects size of these interventions relative to one another over time. Future research is needed to understand the extent to which effects of strategies based on values-oriented, structure-oriented, or dualistic approaches are strong, durable, and persistent.

CONCLUSION

Given the rapid growth within the last two decades in research on the motivations behind and ramifications of unethical behaviors, we argue in this paper that descriptive research on dishonesty also has prescriptive implications in preventing dishonest acts and promoting ethical behaviors. Although by no means exhaustive, the distinction we drew between values-oriented and structure-oriented approaches demonstrates that there is no one right approach to reduce unethical behavior in organizations and society more broadly. Just as it is may not be practical, feasible, or effective to remind people of their morality before encountering every opportunity to cheat, it is also not possible to eliminate all temptations or opportunities to behave unethically. Thus, both researchers and practitioners can benefit from better understandings of how these interventions that target a variety of unethical behaviors can work in concert to fix our moral bugs.

References

- Ainslie, G. (1992). *Picoeconomics: The strategic interaction of successive motivational states within the person*. Cambridge: Cambridge University Press.
- Aquino, K., & Reed, A., II. (2002). The self-importance of moral identity. *Journal of Personality and Social Psychology*, 83(6), 1423-1440. doi: 10.1037/0022-3514.83.6.1423
- Ariely, D., & Wertenbroch, K. (2002). Procrastination, deadline, and performance: Self-control by precommitment. *Psychological Science*, 13(3), 219-224.
- Ashraf, N., Karlan, D., & Yin, W. (2006). Tying odysseus to the mast: Evidence from a commitment savings product in the philippines. *Quarterly Journal of Economics*, 121(2), 635-672.
- Banaji, M. R., Bazerman, M. H., & Chugh, D. (2003). How (un) ethical are you? *Harvard Business Review*, 81(12), 56-65.
- Banaji, M. R., & Greenwald, A. G. (2013). *Blindspot: Hidden biases of good people*. New York: Delacorte Press.
- Bandura, A. (1999). Moral disengagement in the perpetration of inhumanities. *Personality & Social Psychology Review*, 3(3), 193-209.
- Bandura, A., Barbaranelli, C., Caprara, G. V., & Pastorelli, C. (1996). Mechanisms of moral disengagement in the exercise of moral agency. *Journal of Personality and Social Psychology*, 71(2), 364.
- Bateson, M., Nettle, D., & Roberts, G. (2006). Cues of being watched enhance cooperation in a real-world setting. *Biology Letters*, 2(3), 412-414. doi: 10.1098/rsbl.2006.0509

- Batson, C. D., Thompson, E. R., Seuferling, G., Whitney, H., & Strongman, J. A. (1999). Moral hypocrisy: Appearing moral to oneself without being so. *Journal of Personality and Social Psychology*, 77(3), 525-537. doi: 10.1037/0022-3514.77.3.525
- Bazerman, M. H., & Gino, F. (2012). Behavioral ethics: Toward a deeper understanding of moral judgment and dishonesty. *Annual Review of Law and Social Science*, 8(1), 85-104. doi: 10.1146/annurev-lawsocsci-102811-173815
- Bazerman, M. H., Gino, F., Shu, L. L., & Tsay, C.-J. (2013). The power of the cognition/emotion distinction for morality. *Emotion*.
- Bazerman, M. H., & Moore, D. A. (2012). *Judgment in managerial decision making*. New York: John Wiley & Sons.
- Bazerman, M. H., & Tenbrunsel, A. E. (2011). *Blind spots: Why we fail to do what's right and what to do about it*. Princeton: Princeton University Press.
- Becker, G. S. (1978). *The economic approach to human behavior*. Chicago: University of Chicago Press.
- Beshears, J., Choi, J. J., Laibson, D., Madrian, B. C., & Sakong, J. (2012). Self control and liquidity: How to design a commitment contract.
- Bohnet, I., Bazerman, M. H., & Van Geen, A. (2012). When performance trumps gender bias: Joint versus separate evaluation.
- Boyarsky, B. J., Hall, E. C., Deshpande, N. A., Ros, R. L., Montgomery, R. A., Steinwachs, D. M., & Segev, D. L. (2012). Potential limitations of presumed consent legislation. *Transplantation*, 93(2), 136–140.
- Caruso, E. M., & Gino, F. (2010). Blind ethics: Closing one's eyes polarizes moral judgments and discourages dishonest behavior. *Cognition*, 118(2), 280-285.

- Chugh, D., Bazerman, M. H., & Banaji, M. R. (2005). Bounded ethicality as a psychological barrier to recognizing conflicts of interest. In D. A. Moore, D. M. Cain, G. Lowenstein & M. H. Bazerman (Eds.), *Conflicts of interest: Challenges and solutions in business, law, medicine, and public policy* (pp. 74–95). New York: Cambridge University Press.
- Cleek, M. A., & Leonard, S. L. (1998). Can corporate codes of ethics influence behavior? *Journal of Business Ethics*, 17(6), 619-630. doi: 10.1023/a:1017969921581
- Cohen, T. R., Panter, A. T., Turan, N., Morse, L., & Kim, Y. (in press). Moral character in the workplace. *Journal of Personality and Social Psychology*.
- Colby, A., & Damon, W. (1994). *Some do care*. New York: Simon and Schuster.
- Conway, P., & Peetz, J. (2012). When does feeling moral actually make you a better person? Conceptual abstraction moderates whether past moral deeds motivate consistency or compensatory behavior. *Personality and Social Psychology Bulletin*, 38(7), 907-919. doi: 10.1177/0146167212442394
- Cushman, F., Tsay, C.-J., Greene, J. D., & Bazerman, M. H. (2013). Unpublished data.
- Darley, J. M., & Batson, C. D. (1973). 'From jerusalem to jericho': A study of situational and dispositional variables in helping behavior. *Journal of Personality and Social Psychology*, 27(1), 100-108. doi: 10.1037/h0034449
- Desai, S., & Gino, F. (2012). The return to innocence: Nursery rhymes, soft toys, and everyday morality.
- Diener, E., & Wallbom, M. (1976). Effects of self-awareness on antinormative behavior. *Journal of Research in Personality*, 10(1), 107-111. doi: [http://dx.doi.org/10.1016/0092-6566\(76\)90088-X](http://dx.doi.org/10.1016/0092-6566(76)90088-X)

- 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. doi: 10.1177/0146167210385922
- Epley, N., Caruso, E., & Bazerman, M. H. (2006). When perspective taking increases taking: Reactive egoism in social interaction. *Journal of Personality and Social Psychology*, 91(5), 872-889. doi: 10.1037/0022-3514.91.5.872
- Ernest-Jones, M., Nettle, D., & Bateson, M. (2011). Effects of eye images on everyday cooperative behavior: A field experiment. *Evolution and Human Behavior*, 32(3), 172-178. doi: <http://dx.doi.org/10.1016/j.evolhumbehav.2010.10.006>
- Ford, J. D., Ford, L. W., & D'Amelio, A. (2008). Resistance to change: The rest of the story. *Academy of Management Review*, 33(2), 362-377. doi: 10.5465/amr.2008.31193235
- Gino, F., & Ariely, D. (2012). The dark side of creativity: Original thinkers can be more dishonest. *Journal of Personality and Social Psychology*, 102(3), 445-459. doi: 10.1037/a0026406
- Gino, F., Ayal, S., & Ariely, D. (2009). Contagion and differentiation in unethical behavior: The effect of one bad apple on the barrel. *Psychological Science*, 20(3), 393-398. doi: 10.1111/j.1467-9280.2009.02306.x
- Gino, F., & Desai, S. D. (2012). Memory lane and morality: How childhood memories promote prosocial behavior. *Journal of Personality and Social Psychology*, 102(4), 743-758. doi: 10.1037/a0026565
- 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. doi: 10.1016/j.obhdp.2011.01.006

- Gino, F., & Mogilner, C. (2012). Time, money, and morality: How focusing on time, rather than money, decreases dishonest behavior.
- Gino, F., Moore, D. A., & Bazerman, M. H. (Eds.). (2009). *Social decision making: Social dilemmas, social values, and ethical judgments*. New York, NY: Psychology Press.
- 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, A., Imas, A., Brown, A., Nelson, L. D., & Norton, M. I. (2012). Paying to be nice: Consistency and costly prosocial behavior. *Management Science*, 58(1), 179-187. doi: 10.1287/mnsc.1110.1437
- Greenwald, A. G. (1980). The totalitarian ego: Fabrication and revision of personal history. *American Psychologist*, 35(7), 603-618. doi: 10.1037/0003-066x.35.7.603
- Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review*, 102(1), 4-27. doi: 10.1037/0033-295x.102.1.4
- Griffin, D. W., & Ross, L. (1991). Subjective construal, social inference, and human misunderstanding. *Advances in Experimental Social Psychology*, 24, 319-359.
- Gunia, B. C., Wang, L., Huang, L., Wang, J., & Murnighan, J. K. (2012). Contemplation and conversation: Subtle influences on moral decision making. *Academy of Management Journal*, 55(1), 13-33. doi: 10.5465/amj.2009.0873
- Hunt, A., & Rogers, T. (2013). How long do treatment effects last? Persistence and durability of a descriptive norms intervention's effect on energy conservation.
- Jensen, M. C., & Meckling, W. H. (1976). Theory of the firm: Managerial behavior, agency costs and ownership structure. *Journal of Financial Economics*, 3(4), 305-360.

- Johnson, E. J., & Goldstein, D. (2003). Do defaults save lives? *Science*, 302(5649), 1338-1339.
- Jones, T. M. (1991). Ethical decision making by individuals in organizations: An issue-contingent model. *Academy of Management Review*, 16(2), 366-395.
- Just, D. R., & Wansink, B. (2009). Smarter lunchrooms: Using behavioral economics to improve meal selection. *Choices*, 24(3), 1-7.
- Kern, M. C., & Chugh, D. (2009). Bounded ethicality. *Psychological Science*, 20(3), 378-384.
doi: 10.1111/j.1467-9280.2009.02296.x
- Kliver, J., Frazier, R., & Haidt, J. (2014). Behavioral ethics for homo economicus, homo heuristicus, and homo duplex. *Organizational Behavior and Human Decision Processes*, 123(2), 150-158. doi: <http://dx.doi.org/10.1016/j.obhdp.2013.12.004>
- Laibson, D. (1997). Golden eggs and hyperbolic discounting. *Quarterly Journal of Economics*, 112(2), 443-477. doi: 10.1162/003355397555253
- Loewenstein, G., & Prelec, D. (1992). Anomalies in intertemporal choice: Evidence and an interpretation. *Quarterly Journal of Economics*, 107(2), 573-579.
- 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. doi: 10.1509/jmkr.45.6.633
- Mazar, N., & Zhong, C.-B. (2010). Do green products make us better people? *Psychological Science*, 21(4), 494-498. doi: 10.1177/0956797610363538
- McCabe, D. L., Trevino, L. K., & Butterfield, K. D. (1996). The influence of collegiate and corporate codes of conduct on ethics-related behavior in the workplace. *Business Ethics Quarterly*, 6(4), 461-476.

- Merritt, A. C., Effron, D. A., Fein, S., Savitsky, K. K., Tuller, D. M., & Monin, B. (2012). The strategic pursuit of moral credentials. *Journal of Experimental Social Psychology*, 48(3), 774-777. doi: <http://dx.doi.org/10.1016/j.jesp.2011.12.017>
- Messick, D. M., & Bazerman, M. H. (1996). Ethical leadership and the psychology of decision making. *Sloan Management Review*, 37, 9-22.
- Milgram, S. (1974). *Obedience to authority: An experimental view*. New York: Harper & Row.
- Milkman, K. L., Beshears, J., Choi, J. J., Laibson, D., & Madrian, B. C. (2011). Using implementation intentions prompts to enhance influenza vaccination rates. *Proceedings of the National Academy of Sciences*, 108(26), 10415-10420. doi: 10.1073/pnas.1103170108
- Milkman, K. L., Mazza, M. C., Shu, L. L., Tsay, C.-J., & Bazerman, M. H. (2012). Policy bundling to overcome loss aversion: A method for improving legislative outcomes. *Organizational Behavior and Human Decision Processes*, 117(1), 158-167. doi: <http://dx.doi.org/10.1016/j.obhdp.2011.07.001>
- Milkman, K. L., Rogers, T., & Bazerman, M. H. (2008). Harnessing our inner angels and demons: What we have learned about want/should conflicts and how that knowledge can help us reduce short-sighted decision making. *Perspectives on Psychological Science*, 3(4), 324-338. doi: 10.1111/j.1745-6924.2008.00083.x
- Monin, B., & Jordan, A. H. (2009). The dynamic moral self: A social psychological perspective. In D. Narvaez & D. Lapsley (Eds.), *Personality, identity, and character: Explorations in moral psychology* (pp. 341-354). Cambridge: 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. doi: 10.1037/0022-3514.81.1.33

- Moore, C., & Gino, F. (2013). Ethically adrift: How others pull our moral compass from true north. *Research in Organizational Behavior, Forthcoming*.
- Nickerson, D. W., & Rogers, T. (2010). Do you have a voting plan?: Implementation intentions, voter turnout, and organic plan making. *Psychological Science, 21*(2), 194-199. doi: 10.1177/0956797609359326
- Nisan, M. (1991). The moral balance model: Theory and research extending our understanding of moral choice and deviation. *Handbook of Moral Behavior and Development, 3*, 213-249.
- O'Donoghue, T. E. D., & Rabin, M. (1999). Doing it now or later. *American Economic Review, 89*(1), 103-124.
- Oliner, S. P., & Oliner, P. M. (1988). *Altruistic personality: Rescuers of jews in nazi europe*. New York: Touchstone.
- Paharia, N., Kassam, K. S., Greene, J. D., & Bazerman, M. H. (2009). Dirty work, clean hands: The moral psychology of indirect agency. [doi: DOI: 10.1016/j.obhdp.2009.03.002]. *Organizational Behavior and Human Decision Processes, 109*(2), 134-141.
- Phelps, E. S., & Pollak, R. A. (1968). On second-best national saving and game-equilibrium growth. *The Review of Economic Studies, 35*(2), 185-199. doi: 10.2307/2296547
- Prendergast, C. (1999). The provision of incentives in firms. *Journal of Economic Literature, 37*(1), 7-63. doi: doi: 10.1257/jel.37.1.7
- Reed, A., II, & Aquino, K. F. (2003). Moral identity and the expanding circle of moral regard toward out-groups. *Journal of Personality and Social Psychology, 84*(6), 1270-1286. doi: 10.1037/0022-3514.84.6.1270

- Reynolds, S. J., & Ceranic, T. L. (2007). The effects of moral judgment and moral identity on moral behavior: An empirical examination of the moral individual. *Journal of Applied Psychology*, 92(6), 1610-1624. doi: 10.1037/0021-9010.92.6.1610
- Rogers, T., & Bazerman, M. H. (2008). Future lock-in: Future implementation increases selection of 'should' choices. *Organizational Behavior and Human Decision Processes*, 106(1), 1-20. doi: <http://dx.doi.org/10.1016/j.obhdp.2007.08.001>
- Ross, M., & Sicoly, F. (1979). Egocentric biases in availability and attribution. *Journal of Personality and Social Psychology*, 37(3), 322-336. doi: 10.1037/0022-3514.37.3.322
- Sanitioso, R., Kunda, Z., & Fong, G. T. (1990). Motivated recruitment of autobiographical memories. *Journal of Personality and Social Psychology*, 59(2), 229-241. doi: 10.1037/0022-3514.59.2.229
- Schweitzer, M. E., Ordóñez, L., & Douma, B. (2004). Goal setting as a motivator of unethical behavior. *Academy of Management Journal*, 47(3), 422-432.
- Shalvi, S., Eldar, O., & Bereby-Meyer, Y. (2012). Honesty requires time (and lack of justifications). *Psychological Science*, 23(10), 1264-1270. doi: 10.1177/0956797612443835
- 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*, 109(38), 15197-15200. doi: 10.1073/pnas.1209746109

- SKAT. (2005). Tax in denmark: An introduction for new citizens. Retrieved from <http://www.skat.dk/getFile.aspx?Id=52797>
- Stevens, B. (2008). Corporate ethical codes: Effective instruments for influencing behavior. *Journal of Business Ethics*, 78(4), 601-609. doi: 10.2307/25075634
- Tajfel, H., & Turner, J. C. (1986). The social identity theory of intergroup behaviour. In S. Worchel & W. G. Austin (Eds.), *Psychology of intergroup relations* (pp. 7-24). Chicago: Nelson. (Reprinted from: 2nd).
- 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(0), 153-173. doi: <http://dx.doi.org/10.1016/j.riob.2010.08.004>
- Tenbrunsel, A. E., & Messick, D. M. (1999). Sanctioning systems, decision frames, and cooperation. *Administrative Science Quarterly*, 44(4), 684-707. doi: 10.2307/2667052
- Thaler, R. H., & Benartzi, S. (2004). Save more tomorrow™: Using behavioral economics to increase employee saving. *Journal of Political Economy*, 112(S1), S164-S187. doi: 10.1086/380085
- Thaler, R. H., & Sunstein, C. R. (2003). Libertarian paternalism. *American Economic Review*, 93(2), 175-179.
- Thaler, R. H., & Sunstein, C. R. (2008). *Nudge: Improving decisions about health, wealth, and happiness*: Yale University Press.
- Treviño, L. K., Weaver, G. R., & Reynolds, S. J. (2006). Behavioral ethics in organizations: A review. *Journal of Management*, 32(6), 951-990. doi: 10.1177/0149206306294258

Trope, Y., & Liberman, N. (2003). Temporal construal. *Psychological Review*, 110(3), 403-421.

doi: 10.1037/0033-295x.110.3.403

Tversky, A., & Kahneman, D. (1991). Loss aversion in riskless choice: A reference-dependent model. *Quarterly Journal of Economics*, 106(4), 1039-1061.

Vallacher, R. R., & Solodky, M. (1979). Objective self-awareness, standards of evaluation, and moral behavior. *Journal of Experimental Social Psychology*, 15(3), 254-262. doi:

[http://dx.doi.org/10.1016/0022-1031\(79\)90036-2](http://dx.doi.org/10.1016/0022-1031(79)90036-2)

Wade-Benzoni, K. A. (1999). Thinking about the future: An intergenerational perspective on the conflict and compatibility between economic and environmental interests. *American Behavioral Scientist*, 42(8), 1393-1405. doi: 10.1177/00027649921954912

Wegner, D. M. (2003). *The illusion of conscious will*. Cambridge: MIT press.

Welsh, D. T., & Ordóñez, L. D. (2014). The dark side of consecutive high performance goals: Linking goal setting, depletion, and unethical behavior. *Organizational Behavior and Human Decision Processes*, 123(2), 79-89. doi:

<http://dx.doi.org/10.1016/j.obhdp.2013.07.006>

Wright, D. B., Hanoteau, F., Parkinson, C., & Tatham, A. (2010). Perceptions about memory reliability and honesty for children of 3 to 18 years old. *Legal & Criminological*

Psychology, 15(2), 195-207. doi: 10.1348/135532508x400347