

# Testing a Social-Cognitive Model of Moral Behavior: The Interactive Influence of Situations and Moral Identity Centrality

Karl Aquino  
University of British Columbia

Dan Freeman  
University of Delaware

Americus Reed II  
University of Pennsylvania

Vivien K. G. Lim  
National University of Singapore

Will Felps  
Erasmus University

This article proposes and tests a social-cognitive framework for examining the joint influence of situational factors and the centrality of moral identity on moral intentions and behaviors. The authors hypothesized that if a situational factor increases the current accessibility of moral identity within the working self-concept, then it strengthens the motivation to act morally. In contrast, if a situational factor decreases the current accessibility of moral identity, then it weakens the motivation to act morally. The authors also expected the influence of situational factors to vary depending on the extent to which moral identity was central to a person's overall self-conception. Hypotheses derived from the framework were tested in 4 studies. The studies used recalling and reading a list of the Ten Commandments (Study 1), writing a story using morally laden terms (Study 4), and the presence of performance-based financial incentives (Studies 2 and 3) as situational factors. Participants' willingness to initiate a cause-related marketing program (Study 1), lie to a job candidate during a salary negotiation (Studies 2 and 3), and contribute to a public good (Study 4) were examined. Results provide strong support for the proposed framework.

*Keywords:* identity, morality, moral identity, social-cognitive theory, prosocial behavior

The belief that some people are moral and others immoral has always been a convenient way of explaining both the barbarism of a tyrant like Saddam Hussein and the otherworldly self-sacrifice of a Mother Teresa. But dividing the people of the world into the

wicked or the virtuous does not fully capture the contingent nature of human morality. Even the most pious person sometimes violates moral standards he or she claims to hold dear, and even the meanest scoundrel sometimes displays acts of kindness and generosity. That we vacillate from abandoning our moral principles in one situation to acting on them with extraordinary will and determination in another is merely to recognize that in the messy, imperfect world of everyday morality, the situations in which we find ourselves can often be decisive in determining the direction toward which our moral compass turns.

The notion that situational factors influence behavior is one of the foundational assumptions of social psychology (Cervone & Shoda, 1999; Mischel, 1968; Ross & Nisbett, 1991), and many empirical studies support the role of situations as determinants of moral conduct. For example, Darley and Batson (1973) showed that seminary students were much less likely to help an individual in need when situational time pressures constrained such prosocial behavior (i.e., when they were asked to "hurry up"). In a study of negotiator behavior, Aquino (1998) found that negotiators were less likely to deceive their partners when ethical norms within their organization were emphasized. More recently, studies have shown that people are especially likely to act in ways that are harmful to others in situations in which such behavior can be rationalized (e.g., Batson, Kobrynowicz, Dinnerstein, Kampf, & Wilson, 1997; Detert, Treviño, & Sweitzer, 2008; Mazar, Amir, & Ariely, 2008). Although some scholars have gone so far as to argue that moral

---

Karl Aquino, Organizational Behaviour and Human Resource Division, University of British Columbia, Vancouver, British Columbia, Canada; Dan Freeman, Department of Business Administration, University of Delaware; Americus Reed II, Department of Marketing, The Wharton School, University of Pennsylvania; Vivien Lim, Department of Management & Organization, National University of Singapore, Singapore; Will Felps, Department of Organization and Personnel Management, Rotterdam School of Management, Erasmus University, Rotterdam, the Netherlands.

This research was partially supported by a University of Delaware College of Business Administration Grant and a Social Sciences and Humanities Research Grant of Canada awarded to Karl Aquino. Karl Aquino and Dan Freeman contributed equally to this article. We express our gratitude for the helpful editorial guidance provided by Tim Kasser and the thoughtful, constructive feedback we received from reviewers. We would also like to acknowledge Scott Ray, Charlie Field, Mike Gordon, and Phil Oyerly for their assistance with the data collection and coding for Study 3.

Correspondence concerning this article should be addressed to Karl Aquino, Organizational Behaviour and Human Resource Division, 2053 Main Mall, Sauder School of Business, Vancouver, BC V6T 1Z2, or to Dan Freeman, Department of Business Administration, 117 Alfred Lerner Hall, University of Delaware, Newark, DE 19716. E-mail: karl.aquino@sauder.ubc.ca or freemand@udel.edu

behavior is totally determined by the situation (Doris, 1998; Harman, 2003), a substantial body of empirical research shows that stable individual characteristics also matter. Midlarsky, Fagin Jones, and Corley (2005) found that holocaust rescuers were distinguishable from a comparison group of bystanders on the basis of personality variables, including internalized altruistic values, a sense of social responsibility, and empathic concern. In another study, Walker and Frimer (2007) found that personality-related themes like agency and communion were more prevalent in the life narratives of moral exemplars—people who have been recognized for their moral actions—as compared against a demographically matched comparison group.

One interesting aspect of Walker and Frimer's (2007) findings relates to a divergence in the personality profiles of different types of moral exemplars. They found that people who have shown extraordinary and long-term commitment to caring for others (care exemplars) were more nurturant, generative, and optimistic than individuals who have risked their own lives to save others (brave exemplars). To explain this divergence, Walker and Frimer (2007, p. 857) suggest that "powerful situational factors undoubtedly contributed" to the behaviors of brave exemplars. Walker and Frimer's (2007) observations appear to demonstrate the influence of both situational and individual factors on moral behaviors, leading the authors to call for "more systematic exploration of the interaction of personal and situational variables in the moral domain" (p. 857).

In response to their plea, we propose a theoretical framework that explains how situational factors and a relatively stable individual characteristic that we refer to as the *centrality of moral identity* (Aquino & Reed, 2002) jointly influence moral behavior. By *moral behavior* we mean actions that demonstrate social responsiveness to the needs and interests of others, an orientation that many ethicists (e.g., Kant, 1959/1785; Singer, 1981) and psychologists (e.g., Eisenberg, 2000; Gilligan, 1982) view as the defining feature of morality. In the present article, we tested the predictive utility of the framework in four studies. By so doing, we address an unexamined and important gap in present knowledge regarding how situational factors interact with individual characteristics to influence moral functioning (Hardy & Carlo, 2005).

Our theoretical framework is grounded in social-cognitive theory (Bandura, 2001) and uses concepts and psychological mechanisms from social cognition to offer specific predictions about when and how the influence of situational factors on moral behaviors will be moderated by the centrality of moral identity. One premise of the model is that *moral identity*, which we conceptualize as the cognitive schema a person holds about his or her moral character, is a powerful source of moral motivation because people generally desire to maintain self-consistency (Blasi, 1980, 1993, 2004). Thus, someone whose self-definition is organized around moral traits or characteristics should be motivated to behave in a moral manner to maintain this self-conception. A second premise of the model is that people balance multiple facets of their identities, of which only a subset known as the *working self-concept* can be held in consciousness at any given time (Markus & Kunda, 1986; Minsky, 1988). Consequently, the influence of any single facet of identity, including an individual's moral self-conception, will be a function of how accessible that facet of identity is in any given situation (Skitka, 2003). The third and final premise of the model is that situational factors may activate a person's moral

identity or they may activate alternative facets of identity, thereby increasing or decreasing the current accessibility of the moral self-schema within the working self-concept. In combination, these three aspects of the model underlie our overarching hypothesis: Moral intentions and behaviors will be a joint function of (a) the centrality of moral identity to an individual's self-conception and (b) the extent to which situational cues temporarily affect the current accessibility of the moral self-schema within the working self-concept.

### A Social-Cognitive Conception of Moral Identity

A variety of individual characteristics have been proposed as determinants of moral behavior, including moral reasoning (Kohlberg, 1969), moral maturity (Walker & Pitts, 1998), moral commitment (Colby & Damon, 1992), moral personality (Walker & Frimer, 2007), and moral character (Blasi, 2005). Previous research generally finds that these characteristics are associated with moral behavior to varying degrees, but none appears fully capable of predicting or accounting for situational variability in moral behavior (Hardy & Carlo, 2005; Shao, Aquino, & Freeman, 2008). Several authors (cf. Aquino & Reed, in press; Aquino & Reed, 2002; Lapsley, 1996; Lapsley & Narvaez, 2004; Shao et al., 2008; Weaver, 2006) have suggested that social-cognitive theory (Bandura, 2001) may provide a useful framework for addressing this limitation. Emerging research supports this claim by demonstrating the utility of social-cognitive conceptions of moral identity for predicting moral outcomes (Aquino & Reed, 2002; Aquino, Reed, Thau, & Freeman, 2007; Detert et al., 2008; Reed & Aquino, 2003; Reed, Aquino, & Levy, 2007; Reynolds & Ceranic, 2007). On the basis of this evidence, we believe there are strong empirical reasons for adopting a social-cognitive conception of the centrality of moral identity as the cornerstone of our theoretical framework.

From a social-cognitive perspective, a person's moral identity is stored in memory as a complex knowledge structure consisting of moral values, goals, traits, and behavioral scripts (Aquino & Freeman, in press; Aquino & Reed, 2002; Aquino, Reed, Stewart, & Shapiro, 2005; Lapsley & Narvaez, 2004). Because knowledge structures are acquired through life experiences that vary across persons, the importance or centrality of this moral self-schema to one's overall self-conception also differs across individuals (Aquino & Reed, 2002; Blasi, 1980, 2004; Lapsley, 1996; Lapsley & Narvaez, 2004; Narvaez, Lapsley, Hagele, & Lasky, 2006). People whose moral identity occupies greater centrality within the self-concept should perceive that being a moral person is more self-definitional relative to other identities (Blasi, 2004). Consequently, moral identity should exert a stronger influence on processes that guide one's cognition and behavior than other aspects of identity. Furthermore, based on the principle that "neurons that fire together, wire together" (Kihlstrom & Klein, 1994; Minsky, 1988), the moral self-schema of someone for whom moral identity is highly central should be activated more strongly and more frequently than the other self-schemas comprising his or her network of self-identities (cf. Higgins, 1989).

The conception of moral identity centrality we are advocating is closely related to concepts like "schematicity" (see Bem, 1981; Markus, 1977) and "strength of identification" (Deshpande, Hoyer, & Donthu, 1986) that have been used to describe the degree to which a person adopts a particular identity as a basis for his or her

self-definition. Hence, we propose that the greater the centrality of moral identity, the higher its *activation potential* (see Higgins, 1996) and the stronger its ability to affect information processing and moral behavior. *Activation potential* refers to the extent to which a knowledge structure tends to be readily accessible for processing and acting on information (Higgins & Brendl, 1995). The activation potential of any particular knowledge structure can be the result of several factors ranging from frequent priming, to inherited personality orientations, to repeated acts of self-definition (Higgins, 1989).

When the moral self-schema is activated within an individual's working self-concept, we assume that it has a greater potential to influence moral behavior. However, the accessibility of the moral self-schema does not fully explain the motivational impetus provided by moral identity; the basic human desire to maintain self-consistency must also be invoked. According to Blasi (1980, 1993, 2004), the felt obligation to engage in a moral action is related to moral identity through the desire to maintain self-consistency. A person who has a highly central moral identity should feel obligated to adhere to the behavioral prescriptions associated with his or her moral self-schema to avoid self-condemnation. In contrast, a person for whom moral identity is less central to his or her self-concept should not feel any such obligation, and can therefore be expected to be less motivated to engage in moral behavior, meaning that he or she will be less responsive to the needs and interests of others.

### Moral Identity and the Multifaceted Self-Concept

In everyday life, people assume many roles and follow different behavioral scripts across situations. For example, over the course of a few hours, a reader of this article may shift back and forth between being a teacher or scholar, friend or colleague, parent or spouse, life partner or lover. As these roles change, different facets of identity can become more or less salient. The social-cognitive view of the self-concept as a network of identity schemas recognizes that people balance multiple identities and that only a few can be held in consciousness at any given time (Carver & Scheier, 1998; Markus & Kunda, 1986; Minsky, 1988; Skitka, 2003). According to social-cognitive principles, the influence of any identity that comprises the working self-concept is a function of how accessible that identity is in a given situation. As such, defining oneself as a moral person will only produce moral motivation when moral identity is currently accessible (i.e., active) within the working self-concept (Skitka, 2003). Similarly, when a different facet of identity is accessible, people should be more motivated to behave in a manner that is consistent with the values and goals associated with that identity. By specifying the set of circumstances under which a person's moral identity is most likely to be accessible, as well as the set of circumstances under which aspects of identity with inherently oppositional values and goals will be accessible, we can apply our social-cognitive framework to derive specific predictions about the interplay between situational factors and the centrality of moral identity.

### Situational Activation and Deactivation of Moral Identity

We draw from circumplex models of human goals (Grouzet et al., 2005) and values (Schwartz, 1992, 1994; Schwartz & Boehnke,

2004) to provide a theoretical basis for specifying when moral identity will be active or inactive within the working self-concept. The central notion underlying circumplex models of values is that human goals and values can be arrayed in a manner that reflects their degree of compatibility or conflict. In Grouzet et al.'s (2005) model, goals like popularity, image, and financial success are extrinsic and highly compatible with one another. In contrast, goals like affiliation and community feeling are intrinsic and therefore diametrically opposed to extrinsic goals. Similarly, in Schwartz's (1992) model, achievement and power are highly compatible values reflecting self-enhancement. These values are diametrically opposed to universalism and benevolence, which are compatible values reflecting self-transcendence. Both Grouzet et al.'s (2005) and Schwartz's (1992) models have been rigorously validated across diverse cultures (Grouzet et al., 2005; Schwartz, 1992, 1994; Schwartz & Boehnke, 2004), indicating that they are broadly applicable.

Although there are some differences between the two models, they share one important commonality: Both posit an inherently antagonistic relationship between self-transcendent, moral goals and values and self-interested/self-enhancement-related goals and values. This relationship is represented by the diametric opposition (i.e., 180° of separation) between these different types of goals/values in the respective models. For example, in the model of human goals (Grouzet et al., 2005), the goal of being a moral person (called "community feeling" by these scholars and measured with items like "I will assist people who need it and ask nothing in return") was 192° away from the self-interested goal of "financial success," which was measured with items like "I will be financially successful." Similarly, Schwartz and Boehnke (2004) found that the moral value of benevolence, which includes related concepts such as being honest, forgiving, and helpful, are almost exactly in opposition to the self-interested value of personal achievement, which includes related concepts such as being successful and ambitious.

Research aimed at examining how the activation of self-interest-related goals and values affects behavior illustrates the incompatibility between self-interest or financial achievement goals and prosocial or benevolent goals (Kasser, Cohn, Kanner, & Ryan, 2007). For example, Vohs, Mead, and Goode (2006) showed that when reminded of money, participants tend to behave in more selfish and less generous ways. In another study, Kay and his colleagues (Kay, Wheeler, Bargh, & Ross, 2004) showed that presenting people with material objects common to the domain of business (e.g., boardroom, tables, briefcases) increased the cognitive accessibility of the construct of competition and led people to behave less cooperatively when dividing resources and also to interpret ambiguous social interactions as competitive. These findings suggest that when people are primed to focus on material goals or financial success, they are more likely to think about and act in ways that advance their own interests even if it may come at the expense of others.

The inherent incompatibility of moral goals/values and self-interested goals/values suggests that the simultaneous activation of moral identity and a self-interested facet of identity within the working self-concept can produce a dissonant psychological state (see Burroughs & Rindfleisch, 2002). One way to alleviate this aversive state would be for a person to deactivate one of the incompatible facets of identity (Burke, 2003). When one facet of

identity is activated by a situational factor while the other is accessible within the working self-concept due to its having high centrality, we might expect the situation-activated aspect of identity to “win out” most of the time because of its recency of activation and continual reinforcement (assuming the situation remains mostly unchanged). If so, then the situational activation of a self-interested facet of identity may temporarily reduce the accessibility of moral identity for people for whom the latter is highly central. As a result, the psychological tension created by holding conflicting goals/values in consciousness is diminished. Consistent with the social-cognitive principles just discussed, a reduction in the current accessibility of moral identity would weaken its influence on behavior. Similarly, when situational factors activate the moral self-schema, self-interested facets of identity may become less accessible for people for whom such identities are highly central, thereby weakening the influence on behavior of identities that motivate selfishness. As the following sections detail, this theoretical logic provides us with the basis for specifying when and how situational factors will interact with moral identity centrality to influence moral functioning.

### The Interaction of Situational Factors and Moral Identity

The social-cognitive conception of moral identity centrality, the multifaceted self-concept, and the circumplex models of human goals and values can be integrated to derive several specific hypotheses. First, situational factors that activate (or prime) the moral self-schema should increase its accessibility within the working self-concept. Second, because moral identity will tend to be active within the working self-concepts of people for whom moral identity has high centrality, the effect of a situational factor that activates the moral self-schema should be more pronounced (stronger) for people whose moral identity has relatively low centrality. In other words, the prime is expected to activate an otherwise dormant aspect of the self-concept for people low in moral identity centrality but have little influence on people with highly central moral identities because their moral self-schemas are already accessible. Third, the current accessibility of the moral self-schema within the working self-concept should be positively related to moral intentions and behaviors. If these hypotheses are correct, then we would expect to find that the influence of a moral prime on moral intentions and behaviors is moderated by the centrality of moral identity and mediated by the current accessibility of the moral self-schema within the working self-concept. In Studies 1 and 4, we used moral priming manipulations to test these predictions.

Additional hypotheses pertaining to the influence of situational factors that activate self-interested identities can also be derived from our social-cognitive framework. First, situational factors that activate (or prime) a self-interested facet of identity should increase the accessibility of this type of identity, thereby decreasing the accessibility of the moral self-schema. Second, because the activation of a self-interested facet of identity will result in psychological tension if moral identity is also active within the working self-concept, the effect of such a situational factor should be most pronounced among people for whom moral identity has high centrality. In other words, the current accessibility of the moral self-schema within the working self-concept should be an interactive function of the self-interest-promoting situational factor and

moral identity centrality. Third, the current accessibility of the moral self-schema within the working self-concept should be negatively related to intentions and behaviors that can harm others while benefiting the self. If these hypotheses are correct, then we would expect to find that the influence of a self-interest-promoting situational factor on morally questionable intentions and behaviors is moderated by moral identity centrality and mediated by the current accessibility of the moral self-schema within the working self-concept. In Studies 2, 3, and 4, we used self-interest-promoting task manipulations to test these predictions.

### Study 1

To provide an initial test of the hypotheses derived from our social-cognitive model, we examined the joint influence of a moral prime and the centrality of moral identity on participants' intentions to enact a moral behavior. In this study, we examined how trying to recall and then reviewing a list of the Ten Commandments influenced participants' willingness to initiate a cause-related marketing program that would benefit others at a personal cost. According to our model, the moral prime (i.e., the Ten Commandments) should increase intentions to behave morally by increasing the current accessibility of moral identity within the working self-concept. However, the effect of the moral prime should not be uniform for all participants. Rather, it should have the strongest impact on people for whom moral identity is relatively low in centrality because their moral identities are unlikely to be accessible within their working self-concepts in the absence of the prime. Its impact should be weaker, however, for people with highly central moral identities because this identity is already accessible within their working self-concepts.

### Method

#### *Sample and Procedure*

Ninety-two undergraduate business students from the University of Delaware participated for course credit. Participants were recruited for a study of “individual characteristics and decision making.” Fifty-four percent of the participants were men. The sample was homogeneous in terms of age ( $M = 20.3$ ,  $SD = 0.8$ ).

The study consisted of two ostensibly unrelated parts: an online survey and an in-lab priming experiment. The online survey was administered at least 24 hr prior to participation in the experiment. The survey contained a measure of the centrality of moral identity as well as demographic items and several other measures. A two-group design with a control condition and a moral prime condition was used in the in-lab priming experiment. Both groups were first informed that “completion of this survey involves answering some questions about your general knowledge and personal opinions and then making a series of decisions.” They were then asked to complete study tasks in the following order: (a) general knowledge item(s), (b) a measure of the current accessibility of moral identity within the working self-concept, (c) a measure of intention to enact a moral behavior, and (d) demographic items. Participants were randomly assigned to experimental conditions.

#### *Moral Priming Manipulation*

The only difference between the control condition and the moral prime condition involved the general knowledge items. Partici-

pants in the control condition completed just one such item: "Please list the 5 largest cities in the United States." Participants in the moral prime condition also completed a second item: "Please list as many of the 10 commandments as you can." Each general knowledge item was followed by the brief visual presentation of a list of correct responses. The self-generated recall and subsequent viewing of the Ten Commandments served as the moral priming manipulation (cf. Mazar et al., 2008). The use of the Ten Commandments as a moral prime was based on the notion that thinking about the moral principles associated with the commandments should activate morally relevant knowledge structures in memory, including the moral self-schema, thereby increasing the accessibility of moral identity within the working self-concept.

### Measures

*Centrality of moral identity.* The Internalization subscale of Aquino and Reed's (2002) moral identity measure was used to assess centrality in all studies reported herein. To complete this measure, participants were asked to read a list of nine characteristics that might describe a person (i.e., caring, compassionate, fair, friendly, generous, helpful, hardworking, honest, and kind) and then to visualize "the kind of person who has these characteristics [and] imagine how that person would think, feel, and act." The nine characteristics were shown by Aquino and Reed (2002) to capture lay construals of a moral prototype (i.e., a person who is moral). Of importance, the word *moral* is not used in their instrument. After being asked to think about someone who possesses these traits, participants were presented with the five items: (a) "It would make me feel good to be a person who has these characteristics"; (b) "Being someone who has these characteristics is an important part of who I am"; (c) "I would be ashamed to be a person who had these characteristics"; (d) "Having these characteristics is not really important to me"; and (e) "I strongly desire to have these characteristics." A 7-point Likert-type response scale ranging from 1 (*strongly disagree*) to 7 (*strongly agree*) was used for each of the items. After recoding Items 3 and 4 so that higher values reflected greater centrality of moral identity, the five items were averaged to determine the moral identity centrality score for each participant ( $\alpha = .89$ ).

The contributions of the present studies are dependent on the validity of Aquino and Reed's (2002) measure, so it is necessary to comment briefly on its appropriateness and validity. First, it should be noted that the measure was developed on the basis of a social-cognitive conception of a person's moral identity as being defined by a network of trait associations. Aquino and Reed (2002) argued that due to the social-cognitive phenomenon of spreading activation (Collins & Loftus, 1975) among clustered self-relevant moral traits in memory (cf. Conway & Pleydell-Pearce, 2000; Kihlstrom and Klein, 1994), asking people to bring to mind someone who possesses the nine traits included in their measure would increase the accessibility of other traits around which a person's moral identity is organized. As a result, the centrality of this identity to the self can be assessed.

Second, emerging evidence supports the construct and predictive validity of the Internalization subscale as a measure of moral identity centrality. Aquino and Reed (2002; Reed & Aquino, 2003) showed that the items on the measure (a) are internally consistent, (b) show significant test-retest reliability, and (c) have a stable

factor structure. Research has also shown that scores on the Internalization measure are indicative of the activation potential of moral identity (i.e., the likelihood that the moral self-schema will be accessible within the working self-concept, thereby influencing task performance). For example, Aquino and Reed (2002, Study 2, p. 1430) found that scores on the measure correlated strongly with response latencies (Greenwald, McGhee, & Schwartz, 1998) that measure the strength of association between moral traits and self-defined target concepts. Aquino and Reed (2002, Study 4, p. 1433) also showed that the measure correlated positively with the moral content of people's spontaneous self-descriptions. In addition, Freeman and Aquino (2008) found that scores on the measure correlated positively with participants' inclusion of moral traits in responding to the prompt "Who am I?" and also with the number of word stems participants completed with moral rather than nonmoral traits. Together, these findings suggest that higher scores on the Internalization measure are indicative of greater potential for moral identity to be activated within the working self-concept.

Research also supports the predictive validity of the Internalization measure (cf. Aquino et al., 2007; Reed et al., 2007; Reynolds & Ceranic, 2007). For example, participants who scored higher on the measure were more likely to donate to a food drive (Aquino & Reed, 2002) and give money to a charitable organization that benefits an outgroup (Reed & Aquino, 2003). Sage, Kavussanu, and Duda (2006) used a sample of adult male footballers from the United Kingdom to examine the influence of moral identity on behaviors enacted while playing football. Results show a negative relationship between the Internalization measure and antisocial behaviors (e.g., trying to get an opponent injured, diving to fool the referee, and elbowing an opposition player).

*Accessibility of moral identity within the working self-concept.* Valid assessment of the current accessibility of moral identity within the working self-concept required the use of a measure that would not in itself affect the degree to which participants' moral identities were activated. To meet this requirement, a ranking procedure was devised in which participants were asked to "rank the items listed on this screen in terms of who you are at the present moment" from "1 = most reflect how you see yourself" to "5 = least reflect how you see yourself." There are countless identities that study participants could have used as a basis for self-definition in the context of the experimental task, but to make the ranking task manageable, the number of identities to be ranked was limited to five. The moral identity option was "a moral person." The four additional identities were "a successful person," "a family member," "an independent person," and "a student." These alternatives were chosen to be relevant for the student sample while not having any obvious conceptual overlap with moral identity. Options were displayed in random order for each participant. The identity of primary interest was the ranking of the "a moral person" option. For ease of interpretation, rankings of this option were coded such that higher values indicate greater accessibility of the moral self-schema within the working self-concept relative to other identities.

*Intention to enact a moral behavior.* A business-related moral choice was used to create a context in which norms that favor self-interest are likely to be considered acceptable. Administration of the intended choice measure involved presenting participants with the following scenario:

Please imagine that you are the brand manager for a breakfast cereal company. Recently, you were approached by the American Cancer Society (ACS) to initiate a cause-related marketing program. Specifically, ACS would like you to donate 25 cents to a special fund for cancer prevention each time one of your products is purchased. According to your research department, adoption of the program is likely to cost more than it earns through an incremental sales increase. Consequently, IF YOU CHOOSE TO INITIATE THE PROGRAM, YOU WOULD BE LESS LIKELY TO EARN A YEAR-END BONUS.

Participants were then asked to complete two items: (a) "What is the percentage chance that you would choose to initiate the cause-related marketing program? (0 to 100%)" and (b) "How likely are you to initiate the cause-related program?" (1 = *extremely unlikely*, 9 = *extremely likely*). Responses to these items were standardized and averaged to form a measure of intention to enact a moral behavior ( $\alpha = .83$ ). Note that by choosing to adopt the cause-related marketing campaign, the participant would presumably be willing to sacrifice some personal gain (a year-end bonus) to show responsiveness to a social need.

*Control variable.* Gender was controlled for because it has been proposed that men and women vary in their reasoning about moral dilemmas (Gilligan, 1982), with women being more oriented toward an ethic of care. If so, then they might be more motivated to behave in ways that take others' interests into account. Gender was dummy coded (0 = male, 1 = female).

### Analysis

Study 1 was designed to test a mediated moderation hypothesis wherein the centrality of moral identity would moderate the influence of a moral prime on the current accessibility of moral identity within the working self-concept, which would in turn influence moral intentions. Therefore, the present analysis followed established prescriptions for testing mediated moderation hypotheses (see Muller, Judd, & Yzerbyt, 2005). Specifically, the priming manipulation was treated as an independent variable with contrast-coded values (i.e.,  $-1$  and  $+1$ ), the centrality of moral identity was treated as a moderator, and the current accessibility of moral

identity was treated as a mediator. The centrality and current accessibility of moral identity were centered around their means prior to conducting the analysis.

According to Muller et al. (2005), establishing mediated moderation requires estimating parameters for three statistical models. For present purposes, Model 1 involved regressing gender, priming condition, moral identity centrality, and a Priming  $\times$  Centrality interaction term on intentions to initiate a cause-related marketing program. Model 2 involved regressing the same predictors on the current accessibility of moral identity. Model 3 involved regressing gender, priming condition, centrality of moral identity, a Priming  $\times$  Centrality interaction term, current accessibility of moral identity, and a Centrality  $\times$  Current Accessibility interaction term on intentions to initiate a cause-related marketing program. Although it was not pertinent to study hypotheses, the Centrality  $\times$  Current Accessibility interaction term was included in Model 3 to test for an alternative form of mediated moderation (see Muller et al., 2005).

Mediated moderation is indicated if the model estimation results meet four criteria: (a) Model 1 shows a significant effect of the Priming  $\times$  Centrality of Moral Identity interaction on intentions, (b) Model 2 shows a significant effect of the Priming  $\times$  Centrality of Moral Identity interaction on current accessibility, (c) Model 3 shows a significant effect of current accessibility on intentions, and (d) the beta coefficient for the Priming  $\times$  Centrality of Moral Identity interaction estimated in Model 3 is reduced in magnitude (or rendered nonsignificant) in comparison with the same coefficient estimated in Model 1 (Muller et al., 2005).

### Results

As Table 1 illustrates, Muller et al.'s (2005) four criteria for establishing mediated moderation were met. We followed the recommendations of Muller et al. (2005) by calculating the simple effects of experimental condition on the mediator (current accessibility of moral identity within the working self-concept) at  $\pm 1$  standard deviations of the moderator (centrality of moral identity) to examine the nature of the mediated moderation effect. We also

Table 1  
Study 1: Model Estimation Results for Assessing Mediated Moderation of the Moral Priming Effect on Intentions to Initiate a CRM Program

Predictor	Model 1: (DV: Intention to initiate CRM program)		Model 2: (DV: Current accessibility of moral identity)		Model 3: (DV: Intention to initiate CRM program)	
	B	t(87)	B	t(87)	B	t(85)
Gender	.09	0.42	.55	2.41*	-.06	-0.27
Moral prime (PRIME)	.04	.41	.22	1.98†	-.02	-0.24
Centrality of moral identity (MI)	.20	1.67†	.25	1.82†	.13	1.01
Prime $\times$ MI	-.26	-2.08*	-.33	-2.39*	-.14	-0.96
Accessibility of moral identity (ACCESS)					.26	2.82**
MI $\times$ ACCESS					-.04	-0.29
Model R <sup>2</sup>		.09†		.19**		.17*

Note. DV = Dependent Variable; CRM = cause-related marketing.  
†  $p < .10$ . \*  $p < .05$ . \*\*  $p < .01$ .

calculated the simple effect of the mediator on the dependent variable at  $\pm 1$  standard deviations of the moderator.

To calculate the simple effects of the moral prime on the current accessibility of moral identity, we used the regression coefficients from Model 2 for the effect of the moral prime ( $B = .22$ ) and the Prime  $\times$  Moral Identity Centrality interaction ( $B = -.33$ ). The simple effects at one standard deviation (.8095) above and below the mean centrality score were as follows:

For high moral identity centrality (+1 *SD*):  $.22 + (-.33)(.8095) = -.047$ .

For low moral identity centrality (-1 *SD*):  $.22 + (-.33)(-.8095) = .487$ .

To calculate the simple effects of the current accessibility of moral identity on intentions to initiate a cause-related marketing program, we used the regression coefficients from Model 3 for the effect of accessibility ( $B = .26$ ) and the Moral Identity Centrality  $\times$  Accessibility interaction ( $B = -.04$ ). The simple effects at one standard deviation (.8095) above and below the mean centrality score were as follows:

For high moral identity centrality (+1 *SD*):  $.26 + (-.04)(.8095) = .228$ .

For low moral identity centrality (-1 *SD*):  $.26 + (-.04)(-.8095) = .292$ .

We then computed the products of these simple effects at each value of the moderator to assess the total indirect effect of the moral prime through the mediator (see Muller et al., 2005). Results showed that for those lower in moral identity centrality (-1 *SD*), the moral prime had a positive effect ( $.487 \times .292 = .14$ ) on their intentions to enact a moral behavior that worked through an increase in the accessibility of moral identity within the working self-concept. In contrast, the moral prime had a negligible effect on those who rated themselves as being relatively high in the centrality of moral identity ( $-.047 \times .228 = -.01$ ).

### Discussion

Study 1 results support our theorizing. Findings show that when a situational factor (i.e., a moral prime) activates a person's moral self-schema, the likelihood that he or she will intend to behave in a prosocial manner increases. However, the influence of moral priming is not uniform across people. Rather, moral priming appears to have stronger effects on people for whom moral identity has relatively low centrality because such priming has greater potential to increase the accessibility of moral identity within the working self-concept. Conversely, moral priming appears to have weaker effects on people for whom moral identity has relatively high centrality because such priming has lesser potential to increase the accessibility of moral identity within the working self-concept. According to our social-cognitive perspective, the degree to which the moral self-schema is accessible within the working self-concept ultimately determines the extent to which moral identity influences moral outcomes. The observation that the current accessibility of moral identity mediated the interactive influence of the moral prime and the centrality of moral identity on intentions to behave morally supports this notion.

### Study 2

Study 2 was designed to complement Study 1 by examining how a situational factor that activates self-interested facets of identity would influence participants' behavioral intentions. As noted previously, thinking about money (Vohs et al., 2006) and aspects of the business domain (Kay et al., 2004) can promote selfish behavior. On the basis of such findings, we used a financial incentive manipulation in a business context in Study 2 to examine the influence of a self-interest-promoting situational factor on intentions to behave in a manner that harms others and violates a moral standard of honesty. Specifically, we examined how the presence of a financial incentive for negotiating the lowest possible starting salary with a job candidate, which could be facilitated by lying, influenced participants' intentions to lie during the salary negotiation.

Recall that research by Schwartz (1992, 1994; Schwartz & Boehnke, 2004) and Grouzet et al. (2005) shows that self-interested and self-achievement-oriented values and goals are inherently oppositional to self-transcendent moral values and goals. Consequently, the simultaneous activation of moral identity and self-interested facets of identity within the working self-concept can be expected to create a dissonant psychological state (Burroughs & Rindfleisch, 2002), resulting in the temporary deactivation of moral identity in the working self-concept. On the basis of this logic, our hypothesis for Study 2 is one of moderated mediation. Specifically, we hypothesize that the presence of a financial incentive for negotiating the lowest possible starting salary with a job candidate will increase participants' intentions to lie to the candidate by decreasing the current accessibility of moral identity within the working self-concept. However, the effect of the financial incentive should not be uniform for all participants. Rather, it should have a stronger impact on those for whom moral identity is relatively high in centrality because the accessibility of their moral identities can be reduced by the financial incentive. The effect of the financial incentive should be weaker, though, for participants with lower moral identity centrality because their moral identities are already likely to be dormant within the working self-concept even in the absence of a financial incentive.

### Method

#### Sample and Procedure

Fifty-five undergraduate business students from the University of Delaware participated in the study to fulfill a course requirement. Fifty-two percent were men. The average age of participants was 20.7 years ( $SD = 0.9$ ).

Participants were recruited for a "Negotiation Study" and told that they would be asked to fill out two brief surveys prior to completing a 10-min role-playing negotiation. In actuality, participants completed the two surveys only; the role play was a ruse that was used to justify the collection of behavioral intention measures. The first survey involved a variety of individual-difference measures, including a measure of the centrality of moral identity. After completing the first survey, participants were given a set of "confidential instructions" to brief them on the negotiation scenario. These instructions included a statement about a \$100 financial incentive that would be awarded to study participants on the basis of either their performance during the negotiation (finan-

cial incentive, experimental condition) or a random drawing (random incentive, control condition). After participants familiarized themselves with the instructions, they were asked to complete the second survey—a prenegotiation questionnaire—that included a measure of the current accessibility of both moral and achievement-oriented identities, as well as a measure of intentions to lie during the negotiation. When participants completed the prenegotiation questionnaire, they were directed to an adjacent room (ostensibly to begin negotiating), debriefed, and thanked for participating.

*Negotiation scenario.* Role instructions for all participants described the negotiation scenario as a two-party negotiation in which a manager and a job candidate were meeting to see whether they could agree on a starting salary. To make the negotiation scenario relevant to moral functioning, it was necessary to create a situation in which participants would have the opportunity to act in a deceptive manner. Therefore, all participants were assigned to play the role of a manager who possessed several pieces of information that could be used strategically to deceive the candidate, thereby enhancing the likelihood that the candidate would agree to a lower starting salary. First, the instructions informed them that the job the candidate was applying for would be eliminated in 6 months due to an organizational restructuring and that the candidate *did not have this information*. Second, participants were told that the candidate strongly desires to remain in the *same job* for at least 2 years and will accept a lower starting salary in return for a verbal commitment of job stability. Third, participants were told that there were no other qualified candidates being considered at the present time; the candidate was unaware of the lack of other qualified candidates. Finally, participants were informed that a failure to fill the position quickly with a qualified applicant would negatively impact their own yearly performance review.

*Experimental manipulation.* A two-group experimental design with a performance-based financial incentive (experimental) condition and a random incentive (control) condition was used. The role instructions given to both groups were nearly identical. The only difference pertained to the description of a cash prize that participants could win. In the performance condition, participants were told the following: “Negotiating a low salary can benefit you personally in this experiment because the person in the manager’s role who negotiates the lowest salary at the end of this study will earn a \$100 cash prize.” In the random condition, participants were told the following: “For participating in today’s experiment, you may win a \$100 cash prize. The winner of the prize will be *randomly* selected from among all of the persons who played the role of manager during the experiment.”

Across both conditions, participants were informed that by doing better for themselves, they would decrease their negotiating partner’s chances of earning \$100 because the candidate’s prize was dependent on his or her starting salary. (Again, this aspect of the study was a ruse in that no negotiation was to take place.)

### Measures

*Centrality of moral identity.* Aquino and Reed’s (2002) five-item measure of moral identity internalization was again used to measure the centrality of moral identity ( $\alpha = .79$ ).

*Accessibility of moral identity and achievement-oriented identities.* A five-option ranking item was devised to assess the accessibility of moral identity relative to achievement-oriented

identities that may be pertinent in a negotiation. This item asked participants to rank items in terms of “who you are at the present moment” from “1 = the item that most reflects how you see yourself” to “5 = the item that least reflects how you see yourself.” The moral identity option was “a moral person.” The four additional options were “a polite person,” “a creative person,” “a clever person,” and “a pragmatic person.” Creative and polite were chosen to represent stimulation and conformity-related aspects of the self-concept, respectively; neither of these aspects of identity was expected to conflict with either achievement-oriented or moral values (Schwartz, 1992) and goals (Grouzet et al., 2005). Clever and pragmatic were chosen to represent self-enhancing, achievement-oriented facets of identity. As such, the values and goals associated with these facets of identity should be inherently antagonistic to the values and goals associated with self-transcendent aspects of identity like moral identity (Grouzet et al., 2005; Schwartz, 1992). To create a general measure of the accessibility of achievement-oriented identity schemas, participants’ rankings of these two options were averaged (Spearman  $\rho = .36$ ,  $p < .05$ ). For ease of interpretation, moral identity and achievement-oriented identity rankings were coded such that higher values indicate a higher degree of accessibility within the working self-concept.

*Intention to lie.* The moral behavior of interest was the intention to lie during the salary negotiation. Lying was chosen as a focal behavior because studies of moral prototypes show that honesty is among the traits that people use most frequently to define moral character (Aquino & Reed, 2002; Lapsley & Laskey, 2001; Walker & Pitts, 1998). Therefore, it seems reasonable to assume that telling a lie is likely to be experienced by most people as violating self-standards of what it means to be a moral person.<sup>1</sup> Participants answered two questions assessing their intention to lie during the negotiation: (a) “What is the percentage chance that you will tell the job candidate that the position is certain to be eliminated in six months if she/he specifically asks about job security?” (0%–100%) and (b) “If the job candidate specifically asks about job security, how likely are you to tell her/him that the position is certain to be eliminated in six months? (1 = extremely unlikely, 9 = extremely likely).” Responses to these items were recoded so that higher scores reflect an increased likelihood of intending to lie to the job candidate. The items were then standardized and averaged to form a measure of intention to deceive ( $\alpha = .95$ ).

*Control variable.* As in Study 1, gender was controlled for (0 = male, 1 = female).

### Analysis

Prior to conducting hypothesis-relevant analyses, the effects of experimental condition on the current accessibility of moral identity and achievement-oriented identities was examined. Study data were then analyzed using the coding and model estimation proce-

<sup>1</sup> There is evidence that people rate (Peterson, Peterson, & Seeto, 1983) and judge (Bussey, 1999) selfishly motivated lies as worse than so-called white lies that are meant to spare others’ feelings or prevent harm. In our studies, we did not focus on these types of lies, but rather we designed our studies such that lies would primarily benefit the person who lies rather than the target of the lie.



dures recommended by Muller et al. (2005) and described previously for Study 1.

### Results

#### Current Accessibility of Moral Identity

Consistent with our predictions, participants in the performance incentive condition exhibited a lower level of current accessibility of moral identity ( $M = 2.91$ ,  $SD = 1.26$ ) than participants in the random incentive condition ( $M = 3.75$ ,  $SD = 1.00$ ),  $F(1, 53) = 7.58$ ,  $p < .01$ . More important, this decreased accessibility appears to have resulted from an increase in the relative accessibility of achievement-oriented facets of identity, as the average ranking of the clever and pragmatic options provided by performance incentive participants ( $M = 2.96$ ,  $SD = 1.12$ ) was significantly higher than the average ranking provided by random incentive participants ( $M = 2.32$ ,  $SD = 0.86$ ),  $F(1, 53) = 5.68$ ,  $p < .05$ . Thus, the presence of a financial incentive for negotiating the lowest possible starting salary with the job candidate appears to have increased the accessibility of achievement-oriented facets of identity while also decreasing the accessibility of moral identity.

#### Intention to Deceive

Model estimation results for the three models prescribed by Muller et al. (2005) for assessing mediated moderation are shown in Table 2. Results show (a) a significant Experimental Condition  $\times$  Moral Identity Centrality interaction on intentions to deceive, (b) a significant Experimental Condition  $\times$  Moral Identity Centrality interaction on the current accessibility of moral identity, (c) a significant effect of the current accessibility of moral identity on intentions to deceive, and (d) a reduction in the magnitude of the Condition  $\times$  Centrality of Moral Identity interaction on intentions to deceive when the current accessibility of moral identity is included in the model. Together, these findings satisfy the criteria for establishing a mediated moderation effect (Muller et al., 2005).

Using the regression coefficients from Model 2, we calculated the simple effects of experimental condition on the mediator (current accessibility of moral identity) at  $\pm 1$  standard deviations of the moderator (i.e., the mean-centered centrality of moral identity) to assess the nature of the mediated moderation effect. We also used regression coefficients from Model 3 to calculate the simple effect of the mediator on the dependent variable (DV; intention to lie) at  $\pm 1$  standard deviations of the moderator. The simple effects of experimental condition on the accessibility of moral identity at one standard deviation (.5968) above and below the mean centrality score were as follows:

$$\text{For high moral identity centrality (+1 SD): } -.40 + (-.60)(.5968) = -.758.$$

$$\text{For low moral identity centrality (-1 SD): } -.40 + (-.60)(-.5968) = -.042.$$

The simple effects of the current accessibility of moral identity on intentions to lie during a salary negotiation (DV) at one standard deviation (.5968) above and below the mean centrality score were as follows:

$$\text{For high moral identity centrality (+1 SD): } -.25 + (-.01)(.5968) = -.256.$$

$$\text{For low moral identity centrality (-1 SD): } -.25 + (-.01)(-.5968) = -.244.$$

We computed the products of these simple effects at each value of the moderator to assess the total indirect effect of the self-interest-promoting financial incentive through the mediator (see Muller et al., 2005). Results showed that for those higher in moral identity centrality (+1 SD), the presence of a financial incentive for negotiating the lowest possible starting salary had a positive effect ( $-.758 \times -.256 = .19$ ) on their intentions to lie that worked through a sizable decrease in the accessibility of moral identity within the working self-concept. In contrast, the presence of a financial incentive had a negligible effect on those who rated

Table 2  
Study 2: Model Estimation Results for Assessing Mediated Moderation of the Financial Incentive Effect on Intentions to Lie

Predictor	Model 1: (DV: Intention to lie)		Model 2: (DV: Current accessibility of moral identity)		Model 3: (DV: Intention to lie)	
	B	t(50)	B	t(50)	B	t(48)
Gender	-.22	-0.85	.36	1.17	-.13	-0.51
Incentive condition (PERFORM)	.20	1.62	-.40	-2.79**	.10	0.72
Centrality of moral identity (MI)	-.27	-1.22	.29	1.10	-.21	-0.75
PERFORM $\times$ MI	.46	2.21*	-.60	-2.47*	.31	1.28
Accessibility of moral identity (ACCESS)					-.25	-2.08*
MI $\times$ ACCESS					-.01	-0.05
Model R <sup>2</sup>		0.20*		.29**		0.26*

Note. DV = Dependent Variable.  
\*  $p < .05$ . \*\*  $p < .01$ .

themselves as being lower in moral identity centrality ( $-.042 \times -.244 = .01$ ). This pattern of results is consistent with our theorizing.

### Discussion

From a social-cognitive perspective, a situational factor should reduce people's motivation to behave prosocially if it decreases the current accessibility of moral identity within the working self-concept. Findings from Study 2 support this notion. The study shows that the presence of a financial incentive for performing well during a negotiation increased the accessibility of achievement-oriented facets of identity and decreased the accessibility of moral identity, especially for those for whom moral identity had relatively higher centrality. Consequently, intentions to lie increased.

The findings from Studies 1 and 2 support our social-cognitive conception of moral identity centrality and our specifications regarding the manner in which moral primes and self-interest-promoting situational factors can be expected to interact with this construct to influence behavior. However, in both studies we examined behavioral intentions rather than actual behaviors. To address this limitation, we conducted Study 3 to examine the joint effects of financial incentives and the centrality of moral identity on whether people actually lied during a negotiation.

### Study 3

The experimental paradigm used in Study 3 was nearly identical to the one used in Study 2. The main difference between the studies is that participants actually negotiated a starting salary with a job candidate in Study 3. The primary hypothesis is the same as in Study 2: Moral identity centrality will moderate the influence of a financial incentive on lying during the negotiation. Specifically, participants with highly central moral identities will be influenced most strongly by the financial incentive, resulting in an increased likelihood that they will lie to the job candidate. The rationale underlying this prediction again invokes the notion that the simultaneous activation of moral identity and self-interested facets of identity within the working self-concept will result in the deactivation of moral identity when the self-interest facet is induced by situation cues. However, in Study 3 we did not measure the current accessibility of moral identity but rely on the findings from Study 2 to infer that deactivation of moral identity is responsible for any observed increase in lying.

Study 3 extends the findings from Study 2 by also testing whether the influence of a highly central moral identity on lying would be sensitive to the magnitude of the incentive and not just its presence. It makes intuitive sense that a small incentive for lying may be insufficient for motivating a person with a highly central moral identity to act dishonestly, but a large incentive might be sufficient. Therefore, we expected that the combination of a performance incentive and receiving a larger, rather than smaller, performance reward would produce the strongest motivation to lie by weakening the ability of a highly central moral identity to discourage lying.

## Method

### Sample and Procedure

Two hundred twenty-four undergraduate business students from the University of Delaware participated in the study to fulfill a course requirement. Forty-two percent were men. The average age of participants was 20.3 years ( $SD = 1.2$ ); their average number of years of work experience was 4.6 ( $SD = 2.0$ ).

A  $2 \times 2$  (Incentive Condition  $\times$  Reward Size) factorial design was used, with the centrality of moral identity as a measured variable. Data were collected in two parts. In the first part, participants completed a personal opinion questionnaire that contained measures of the centrality of moral identity, demographic characteristics, and a variety of other measures. The questionnaire was completed *at least two days prior* to the negotiation. After completing the questionnaire, participants signed up for the second part of the study, which involved a role-playing negotiation.

When participants arrived for their negotiation session, they were randomly assigned to experimental conditions and a role as either a manager or a job candidate and placed in separate rooms. They were then given common background information as well as instructions unique to each role. The instructions and roles were nearly identical to those from Study 2. The one important difference in the instructions pertained to the job candidate's role instructions. To create a situation in which participants playing the manager's role would have an opportunity to lie, the candidate's role instructions made it clear that job stability was an important issue for them and that they should not accept any salary offer unless they received a verbal guarantee from the manager that they would remain at the same job for at least 2 years. The instructions also indicated that the candidates should ask managers a question regarding job stability at the beginning of the negotiation prior to discussing starting salary. This feature of the simulation forced managers to decide whether to lie about a fact that they knew to be false because they knew that the job the candidate was applying for would definitely be eliminated in 6 months.

After the participants read the common and unique role instructions, they were given a prebargaining questionnaire to assess their understanding of the task. The experimenter then asked participants whether they had any final clarification questions. When all questions were answered, participants playing each role were brought together in another room and given 15 min to negotiate. They were told to reach agreement only on the candidate's starting salary and to write the agreement on a contract form. If they failed to agree in 15 min, then they were told that the negotiation reached an impasse and that no future negotiations would take place. In either case, when participants finished the negotiating session, they completed a postexperiment questionnaire. They were then debriefed and thanked for their participation.

### Experimental Manipulations

*Performance incentive.* The conditions for receiving a reward and the size of the rewards were varied for participants in the manager's role as part of the experimental manipulations. The reward was either dependent on the manager's performance or completely independent of performance. Presumably, the incentive to perform well would be stronger in the former than in the latter condition. In the performance condition, managers were told:

Negotiating a low salary can also benefit you personally in this experiment because the person in the manager's role who negotiates the *lowest* salary at the end of this study will earn a (\$150/\$50) cash prize. In the event of a tie, the winner of the prize will be randomly selected.

In the random condition, managers were told: "For participating in today's experiment, you may win a (\$150/\$50) cash prize. The winner of the prize will be *randomly* selected from among all of the persons who played the role of manager during the experiment."

Across all conditions, managers were informed that by doing better for themselves, they would decrease their negotiating partner's chances of earning \$50 because the candidate's prize was dependent on his or her starting salary. The following information was presented to all managers to make them aware that agreeing on a low salary could harm the other party:

You should also know that the person playing the candidate's role today has been told that s/he can earn a \$50 cash prize that s/he can keep if s/he negotiates the highest salary among all the participants who play the same role as him/her during this experiment. So although you and your partner are not directly competing for the same prizes, the procedure used to determine the prize means that negotiating a low salary decreases your partner's chances of earning \$50.

*Reward size.* This manipulation varied the size of the cash reward for people in the manager's role. The prize was \$150 in the "high reward" condition and \$50 in the "low reward" condition.

### Measures

*Lying.* Lying was assessed objectively by videotaping the negotiations and having two master's of business administration research assistants independently code how managers responded to the candidate's question regarding guaranteed job stability. A strict standard was used to code lying in this study. The standard was based on Bok's (1978) definition of lying as deliberately stating something that one knows to be false. In the context of the present study, a statement by the manager providing a verbal guarantee that the candidate would be able to remain on the same job for more than 6 months was coded as a lie. The coders were trained by Karl Aquino to distinguish four possible responses to this question: (a) lying (e.g., "I can guarantee you that you will be at the same job for at least two years"), (b) concealing (e.g., "It's possible you may be on the job for at least two years"), (c) refusing to answer the question (e.g., "I can't tell you that), and (d) telling the truth (e.g., "The job will be restructured after six months"). The coders were naïve to study hypotheses, participants' performance condition, and moral identity scores. After training, the assistants coded the remaining videotapes. The dependent variable was therefore composed of a four-level categorical variable coded 1–4 to represent lying, concealing, refusing to answer, and telling the truth, respectively. The proportional agreement between coders was .89, indicating substantial convergence in their ratings. Disagreements were discussed by the coders and resolved such that both parties agreed on how the manager's response should be classified.

*Centrality of moral identity.* Aquino and Reed's (2002) five-item Moral Identity Internalization scale was used to measure the centrality of moral identity. In contrast to Studies 1 and 2, in which a 7-point response scale was used, respondents answered on a

5-point scale ranging from 1 (*strongly disagree*) to 5 (*strongly agree*). Items were averaged ( $\alpha = .82$ ).

*Control variables.* Gender was controlled for (0 = male, 1 = female).

## Results

### Manipulation Checks

Prior to the negotiation, we assessed managers' understanding of their confidential information by asking them to answer (a) whether the job candidate would be willing to accept a lower salary in exchange for job stability, (b) whether the candidate had any knowledge that the job for which they were being hired would be eliminated in 6 months, and (c) what they would receive if they negotiated the *lowest* salary among all the persons who are assigned to the manager role in this study. All managers answered these questions correctly, increasing confidence that the task was understood, and thus that any deception by those in the managerial role was intentional.

### Hypothesis Tests

One hundred dyads reached agreement and 12 impassed. Reaching agreement was not relevant to our hypotheses, so we analyzed data from all of the dyads. Of the participants, 22% in the manager's role were coded as having lied, 19% as having concealed, 36% as not answering the question, and 23% as telling the truth.<sup>2</sup> The dependent variable consisted of four nonordered categorical responses, so we used multinomial logistic regression with maximum-likelihood estimation to conduct an initial test of our hypotheses. The multinomial model simultaneously estimates the probabilities of behaving one way (e.g., lying) versus another (e.g., telling the truth) and tests whether these probabilities differ as a function of controls, independent variables, and interaction terms. Gender was included as a control variable in the model along with the main effects of moral identity centrality, incentive condition, and reward size. Incentive condition (0 = random, 1 = performance) and reward size (0 = low [\$50], 1 = high [\$150]) were dummy coded. The model also included all two-way interactions involving moral identity centrality, incentive condition, and reward size as well as the three-way Centrality  $\times$  Incentive Condition  $\times$  Reward Size interaction. As in our previous studies, each variable used to construct a multiplicative interaction term was mean centered prior to the computation of the term (Aiken & West, 1991). Lying was treated as the reference category, so the relevant contrasts are between lying and concealing, lying and refusing to answer and lying and telling the truth.

We first tested the fully parameterized model just described. The initial log likelihood value showed improved overall fit after accounting for the independent variables,  $\chi^2(24, N = 112) = 37.05, p < .05$ , indicating a systematic relationship between the behavioral outcomes of the negotiation task and the predictors.

<sup>2</sup> Participants in the manager role who were coded as having lied negotiated more personally favorable agreements ( $M = \$47,776$ ) compared with those who did not lie ( $M = \$51,147; p < .001$ ). This difference indicates that negotiators who lied gained a significant bargaining advantage compared with those who did not lie.

Nagelkerke's pseudo  $R^2$  measure was .30, suggesting that these relationships were moderately strong. Examination of the likelihood ratio tests on the overall dispersion in the pattern of possible responses showed a main effect of moral identity centrality,  $\chi^2(3, N = 112) = 9.47, p < .05$ . However, this main was qualified by a significant Centrality by Incentive interaction,  $\chi^2(3, N = 112) = 7.82, p = .05$ . No other main or interaction effects were significant.

Figure 1 allows us to interpret the nature of the interaction by graphically depicting the percentage of participants who lied, concealed, did not answer, and told the truth as a function of the centrality of moral identity (high vs. low) and the incentive manipulation. Assignment to the moral identity categories was based on a median split on the centrality scores.

Of particular interest for assessing whether Study 3 replicates the findings of Study 2 is the comparison of whether people whose moral identity was either higher or lower in centrality differed in their willingness to lie as a function of the performance incentive manipulation. Figure 1 provides preliminary support for this hypothesis by showing that although the performance incentive manipulation did not appear to influence lying for participants lower in moral identity centrality (i.e., the proportion of those who lied is approximately equal in both conditions), it did so for participants higher in moral identity such that the latter lied more when performance incentives were present rather than absent. We con-

ducted a formal test of whether this observed difference was statistically significant by creating two response categories: lie vs. all other responses. We then performed a logistic regression to predict lying as a function of the performance incentive in the high- ( $N = 57$ ) and low- ( $N = 55$ ) moral identity centrality groups. We also included gender and reward size as predictors in the regression. The analysis showed that participants high in moral identity centrality were more likely to lie in the performance incentive condition compared with the random condition ( $B = 2.02, Wald = 4.90, p < .05$ ). However, the incentive manipulation had no effect on lying for participants who were low in moral identity centrality ( $B = .20, Wald = .11, ns$ ).

### Discussion

Study 3 provides behavioral evidence suggesting that the presence of a self-interest-promoting situational factor like a financial incentive increased people's willingness to lie to another person during an actual negotiation. The pattern of findings is consistent with the specifications of our social-cognitive model as well as with the results of Study 2. If, as Study 2 demonstrated, the presence of a financial incentive decreased the current accessibility of moral identity within the working self-concept, and this effect was more pronounced for participants with higher as opposed to lower moral identity centrality, then we would expect higher moral identity centrality participants to be less motivated to pursue moral values and goals in the performance incentive than in the random condition. In contrast, we would not expect the financial incentive to influence the current accessibility of nonmoral (i.e., self-interested) goals for those whose moral identity already has low centrality, and so their willingness to lie would be unaffected by either the presence or the absence of an incentive. Indeed, this is what our data showed.

We did not find support for the hypothesized three-way interaction involving moral identity, incentives, and reward size. Given our prior results showing the mediating role played by a currently accessible moral identity, one plausible explanation for the null result is that the \$50 incentive was as powerful a situational cue as a \$150 incentive for decreasing the accessibility of moral identity within the working self-concept in our student sample. This explanation does not rule out the possibility that if the difference in reward sizes was more extreme (e.g., \$5 vs. \$500), then a three-way interaction would be found.

The findings of Studies 1, 2, and 3 provide compelling support for our social-cognitive framework. Studies 1 and 2 showed that the current accessibility of moral identity within the working self-concept links situational factors and moral identity centrality to moral action. Study 1 showed that priming people to activate moral constructs increased the accessibility of moral identity within the working self-concept, especially for people for whom moral identity has lower as opposed to higher centrality. Study 2 showed that the presence of a financial incentive for acting in a self-interested way decreased the current accessibility of moral identity relative to inherently conflicting identity schemas, especially for those for whom moral identity has higher as opposed to lower centrality. Moreover, this study showed that such decreased accessibility of the moral self-concept led to an increased willingness to lie in a negotiation. Study 3 suggests that the presence of a financial incentive for acting in a self-interested manner can

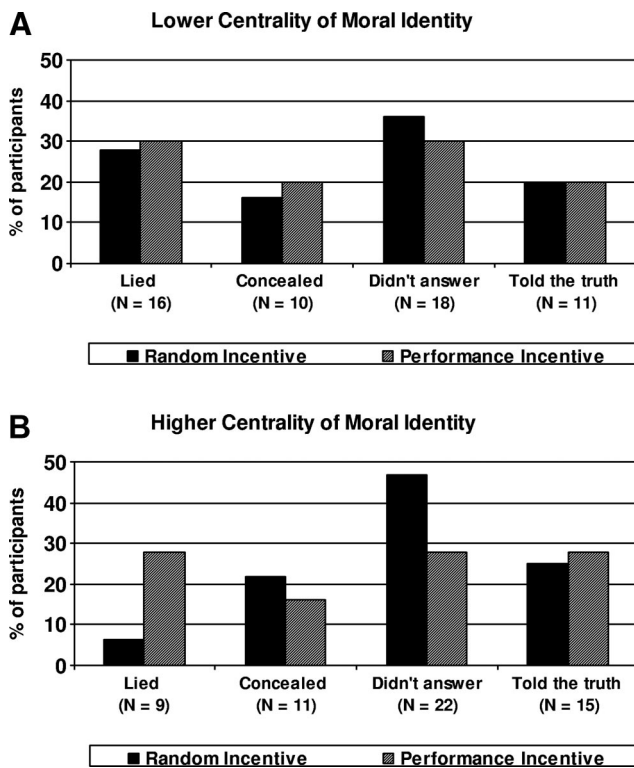


Figure 1. A: Study 3: Percentage of participants who lied, concealed, did not answer, and told the truth by incentive condition for participants with lower moral identity centrality. B: Study 3: Percentage of participants who lied, concealed, did not answer, and told the truth by incentive condition for participants with higher moral identity centrality.

affect actual lying behavior, even for people with relatively high moral identity centrality.

In Study 4, we combine elements of the three prior studies to conduct a final test of our theoretical model. Specifically, we include a moral prime (analogous to Study 1) and present participants with feedback about others' behavior that could activate a more self-interested orientation (analogous to Studies 2 and 3). These design elements allowed us to test whether priming the moral self-schema can help to sustain the influence of a highly central moral identity on prosocial behavior in the face of self-interest-promoting situational cues.

#### Study 4

Study 4 involved a situation that required participants to make a series of choices about whether to contribute to the provision of a public good. Public goods fall within the broader domain of social dilemmas and are a class of situations in which individual rationality conflicts with collective welfare (Komorita & Parks, 1994). The choice inherent in social dilemmas is whether to act solely in one's self-interest or to sacrifice some of one's interests to benefit others. Accordingly, a social dilemma is a highly appropriate situation for exploring the role of moral identity centrality in moral functioning because it pits self-interested motives against the motive to show responsiveness to the interests of others.

Study participants were presented with a particular type of social dilemma that interdependence theorists call a "martyr" situation because a person must sacrifice his or her own self-interest to advance the collective good (Kelley et al., 2003, p. 219). In essence, those who act cooperatively must sacrifice themselves to improve group outcomes. Although not economically rational, these behaviors parallel acts of benevolence and self-sacrifice that demonstrate responsiveness to the needs of others. A norm of self-interest and an inducement to act noncooperatively was emphasized by giving participants false feedback in an iterated game indicating that defection was the most common course of action taken by other members in their group. This procedure created a particularly powerful martyr situation, such that those who continued to choose to cooperate under these circumstances did so knowing that it would harm them and that others have demonstrated selfish behavior.

Study 4 combines elements from prior studies to provide the most complete test of our theoretical arguments. Like Study 1, Study 4 considers the interaction between a moral prime and moral identity centrality in predicting a morally relevant outcome. However, as the experiment progresses through repeated trials, it becomes increasingly apparent that cooperative, self-transcendent behavior leads to lower personal achievement. Thus, like Studies 2 and 3, Study 4 introduces situational cues that should activate achievement-oriented facets of identity.

On the basis of our theoretical framework and the empirical findings of our first three studies, we can make several predictions about what will happen at various stages of the social dilemma. In the initial stage of the social dilemma, we predict that the moral prime should exert less influence on the behavior of people for whom moral identity has high centrality. For such people, the moral self-schema is likely to be resident in the working self-concept, and so the introduction of a prime should have a minimal influence on its current accessibility (see Study 1 findings). How-

ever, the moral prime should initially influence people whose moral identity is less central to contribute more to the public good compared with when they are not primed.

Our second prediction is that over time, as participants receive feedback indicating that others are behaving selfishly, self-interest-oriented facets of identity should become increasingly accessible within the working self-concept, thereby deactivating moral identity, especially for those high in moral identity centrality. Thus, participants' cooperation will decrease in the later stages of the experiment relative to the early stage, a result that would parallel what we found in Studies 2 and 3 in which financial incentives increased intentions to lie and actual lying among those with higher moral identity centrality.

Our third prediction, though, is that the decrease in cooperation over time for those for whom moral identity has high centrality will be less likely to occur when the moral self-schema has been primed. If moral identity has been activated by situational cues, then these cues can act as a situational reinforcement of the current accessibility of moral identity among those for whom moral identity already has high centrality. As a result, such participants should be more motivated to sustain contributions to a public good over time, even when they receive feedback that others are acting selfishly, compared with those who are also high in moral identity centrality but who have not been exposed to the moral prime.

#### Method

##### Sample and Procedure

Thirty-three undergraduate business students from the University of Washington participated in this study to fulfill a course requirement. Sixty-three percent of the participants were men. The sample was ethnically diverse, with 45% identifying themselves as White, 27% as Asian, 18% as Middle Eastern, 7% as African American, and 3% as "other." The average age of participants was 20.1 ( $SD = 2.2$ ).

Data collection was again separated into two parts—a Part 1 survey and a Part 2 experimental session. At least 24 hr prior to their experimental session, participants were required to fill out an online survey that included measures of the centrality of moral identity and a measure of social value orientation (SVO). Participants then came to the behavioral lab for Part 2 of the study. Upon entering the lab, participants were randomly assigned to either a *moral prime* or *control group* condition. Both groups completed a version of what was labeled a "handwriting task." The cover story for this activity was that another professor was interested in "graphology," or the ability to capture personality differences by examining the details of a person's handwriting. Participants were given a  $9 \times 5$  matrix that contained nine words listed in the first column of each row. In the moral prime condition, these words reflected moral traits (e.g., *caring*, *compassionate*, *fair*); in the control condition, these words denoted everyday household objects without moral content (e.g., *book*, *chair*, and the like). Participants in both groups were instructed to write these words (in their normal handwriting) across the remaining four columns such that each participant wrote down the nine words four separate times. On the next page, participants were told to "take a few moments to think about each of these words." They were then instructed to "write a brief story about [themselves] (in one or two paragraphs) which [used] each of these words at least once."

After finishing the handwriting task, participants were sent to a computer terminal to participate in “a virtual task involving investment decisions,” which they were told would be carried out in conjunction with the 4 other participants in the laboratory. Each experimental session involved 5 participants, who together constituted the “group.”<sup>3</sup> Participants were always separated by at least two computers so that others’ actions could not be viewed. Elaborate procedures were used to convince participants that they were receiving feedback about other participants’ choices via a local area network (i.e., LAN line). In reality, each participant received identical feedback manufactured by the experimenter.

Instructions describing the task were placed at each terminal. These instructions indicated that participants would be asked to make a series of investment decisions. On each decision trial, they, as well as every other member of the group, would have to decide whether to allocate 10 points to either a “joint account” or to a “personal account.” The points in the personal account would earn no interest but would not be shared. The points in the joint account would earn 100% interest but would be distributed equally among all group members regardless of their contributions to the joint account. For example, if 3 of the 5 group members contributed to the joint account and the remaining 2 kept their points in the personal account, then the payoff would be 12 for those who contributed to the joint account ( $[10 + 10 + 10] \times 2$  divided among 5 people yields 12 per person) and 22 for those who did not (10 points in the personal account plus the 12 shared points). Thus, the game was structured as a social dilemma in which individual rationality was in conflict with the collective good. In any given iteration, it is advantageous to put resources in the personal account. But if everyone put their resources in personal accounts, then the group misses out on the opportunity to earn interest. Note also that the game was largely framed in economic terms such as “investments,” “points,” and “interest.” As such, the game itself was likely to put people in an “economic” frame of mind, in which self-interest was the presumed norm (Pillutla & Chen, 1999).

Participants were asked to make a series of 20 decisions that were divided a priori into five trial blocks (Pillutla & Chen, 1999). The feedback was manipulated such that during each trial block, defection appeared to be the most common course of action with group members opting for the personal account in 10 of the 16 choices (this set excludes the participant’s own choices) to establish a norm of noncooperation. In a particular trial, this equates to an average of 2.5 of the other 4 members choosing the personal account. This translates to an expected value for contributing to the joint account of just 5 points, which is exactly half of the expected value of contributing to the personal account (10 points). After completing the social dilemma task, participants were asked to write down their guesses about the purpose of the study and to note whether they were suspicious of any part of the experiment. There was no evidence of suspicion of the deceptions (i.e., the graphology cover story and the manufactured feedback), and none of the participants correctly guessed what was being studied. Finally, the participants were debriefed with an explanation of the study’s purpose and its deceptive elements.

### *Moral Priming Manipulation*

As described previously, a handwriting task was used to prime moral identity. This task has been used previously (Aquino et al., 2007; Reed et al., 2007) and shown to activate moral identity successfully, thereby increasing its current accessibility within the working self-concept. The prime manipulation was dummy coded in our analysis (0 = control, 1 = moral prime).

### *Measures*

*Centrality of moral identity.* The five internalization items from Aquino and Reed’s (2002) scale were used to measure the centrality of moral identity. Participants responded on a 5-point scale ranging from 1 (*strongly disagree*) to 5 (*strongly agree*). Items were averaged ( $\alpha = .86$ ).

*Control variables.* Gender was again controlled for (0 = male, 1 = female). SVO, a construct that has been found to be among the most reliable individual-difference predictors of cooperation in many kinds of social dilemmas (Offerman, Sonnemans, & Schram, 1996), was also included as a control variable. Controlling for SVO allowed us to test whether the effects of moral identity centrality would hold even after accounting for a well-established predictor of cooperation. If so, then it would provide strong evidence that the measure of moral identity centrality is not simply an alternative way of measuring social values. SVO was assessed using the RING test (Liebrand & McClintock, 1988). This test asks participants to choose between a series of two different hypothetical payoff allocations for the participant (self) and the hypothetical person (other). For example, a choice might consist of Alternative A (self: \$15, other: \$0), or Alternative B (self: \$13, other: \$7.5). There are 24 choices like this, with the property that across all alternative items, mean (self) = mean (other) = 0. SVO is determined by taking the ArcTan of mean (other) divided by mean (self) and then using that angle to classify people into one of four categories (i.e., altruistic [coded 1] if  $67.5^\circ < \text{angle} < 112.5^\circ$ , cooperative [coded 2] if  $22.5^\circ < \text{angle} < 67.5^\circ$ , individualistic [coded 3] if  $-22.5^\circ < \text{angle} < 22.5^\circ$ , or competitive [coded 4] if  $-67.5^\circ < \text{angle} < -22.5^\circ$ ).

*Cooperation.* The dependent variable of cooperation was calculated by summing the number of times the participant contributed to the joint account (personal = 0, joint = 1) across the four consecutive trials in each trial block. Thus, this was a repeated measure with five observations for each participant; observations ranged in value from 0 to 4.

### *Results*

We performed a repeated measures analysis of covariance to test the effects of the independent variables across the five trial blocks. We entered the prime manipulation as a fixed factor in the model; trial block was the repeated measures factor. We entered gender, SVO, moral identity centrality, and a Centrality  $\times$  Prime Condition interaction into the analysis as covariates. The variables com-

<sup>3</sup> However, 2 participants who had been randomly assigned to the moral prime condition did not follow the instructions to fill out the Part 1 questionnaire prior to participating in Part 2, and thus their data were unusable.

prising the Centrality  $\times$  Prime condition interaction term were mean centered. The analysis included the three-way interaction of centrality, prime condition, and trial block.

Results showed a significant main effect of moral identity centrality,  $F(1, 27) = 9.16, p < .01, \eta^2 = .25$ , and a significant Moral Identity Centrality  $\times$  Moral Prime Condition interaction,  $F(1, 27) = 6.54, p < .05, \eta^2 = .19$ , across trial blocks. We also found a significant Moral Identity Centrality  $\times$  Prime  $\times$  Trial Block interaction,  $F(4, 24) = 2.78, p = .05, \eta^2 = .32$ , and a significant Gender  $\times$  Trial Block interaction,  $F(4, 24) = 3.22, p < .05, \eta^2 = .35$ . We did not propose any hypotheses about the effect of gender on cooperation, nor was the pattern of the interaction involving gender clearly interpretable. Hence, this interaction effect is not discussed further. Instead, we turn our attention to the interaction involving the moral prime and moral identity centrality.

We analyzed the Moral Identity Centrality  $\times$  Moral Prime interaction by summing the number of cooperative choices made across all trials to produce an overall cooperation score. We then regressed overall cooperation on the moral prime manipulation in groups composed of participants who were either high ( $N = 16$ ) or low ( $N = 17$ ) in moral identity centrality. Group assignment was determined by a median split on moral identity centrality. The results of this analysis indicate that the moral prime did not predict overall cooperation among people who were low in moral identity centrality ( $B = 1.09, ns, R^2 = .04$ ), but did so for participants high in moral identity centrality ( $B = 3.88, p < .01, R^2 = .43$ ). These results support our hypothesis that priming the moral self-schema would motivate participants to sustain cooperation over time despite the defection of others, but only if they were high rather than low in moral identity centrality.

We analyzed the three-way interaction between moral identity centrality, moral prime condition, and trial block by first splitting

the sample into high- ( $N = 16$ ) and low- ( $N = 17$ ) moral identity centrality groups and then regressing cooperation in each of the five trial blocks on the dummy-coded moral prime variable. Results showed that in Trial Block 1 (Trials 1–4), participants with lower moral identity centrality who completed the moral priming task cooperated more than comparable participants in the control group ( $B = 1.26, p < .05, R^2 = .30$ ); however, in this first trial block, there was no significant difference in cooperation as a function of the prime for participants higher in moral identity centrality ( $B = -0.13, ns, R^2 = .01$ ). This pattern of results supports our prediction that the moral prime would initially be less influential among those for whom moral identity already has high centrality, which is consistent with what we found in Study 1.

Then, as participants received feedback about the selfish behavior of other group members, the results showed that participants with higher moral identity centrality reacted to the moral prime in a way that sustained their cooperation in the face of others' defection. In Trial Block 4 (Trials 13–16), participants with higher moral identity centrality in the moral prime condition cooperated more than comparable participants in the control condition ( $B = 1.00, p = .06, R^2 = .23$ ); however, the prime condition did not have a significant impact on cooperation for participants with lower centrality ( $B = -0.66, ns, R^2 = .11$ ). In the fifth and final trial block (Trials 17–20), we again found that participants with higher moral identity centrality who had been primed cooperated more than comparable participants in the control group ( $B = 2.00, p < .01, R^2 = .58$ ), but once again there was no significant difference in cooperation as a function of the prime for those with lower moral identity centrality ( $B = -0.03, ns, R^2 = .00$ ).

Figure 2 depicts the mean levels of cooperation across the five trial blocks for participants with either higher or lower moral identity centrality in either the moral prime or control conditions.

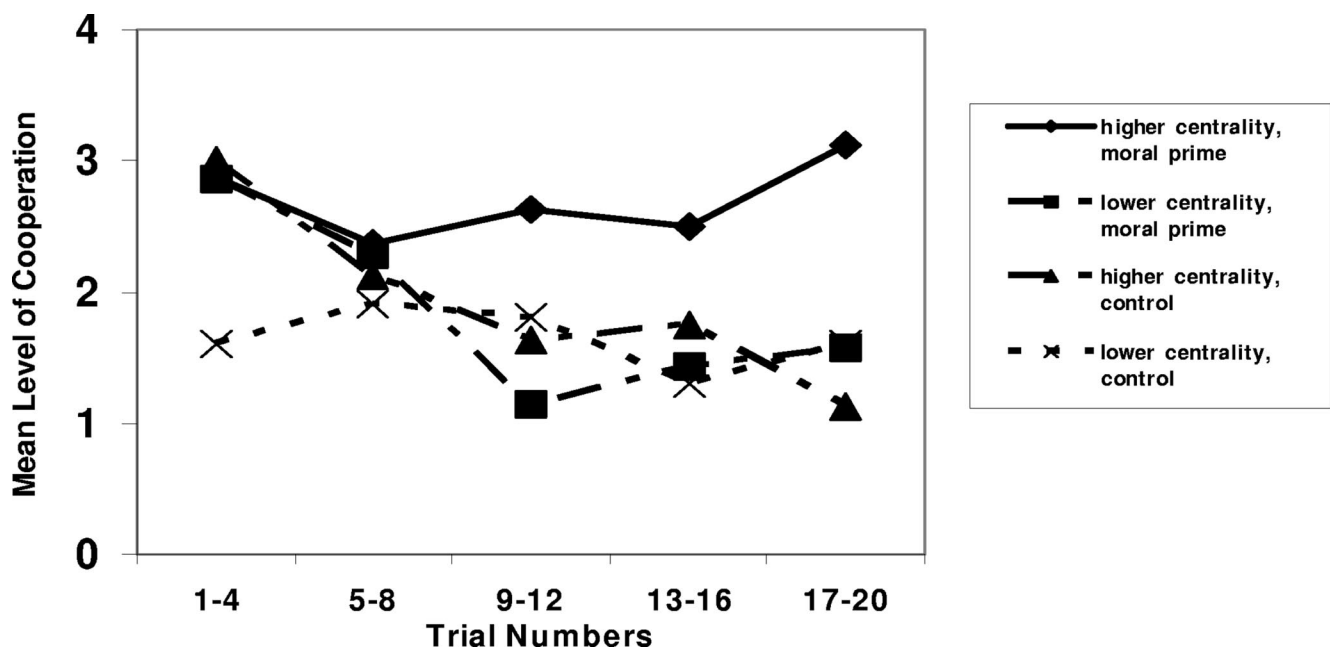


Figure 2. Study 4: Mean level of cooperation by participants higher and lower in moral identity centrality across trial blocks in the moral prime and control conditions.

For illustrative purposes, assignment to higher and lower moral identity groups was based on a median split on the centrality measure. The pattern shows that higher centrality participants who received the moral prime sustained higher levels of cooperation in the face of others' selfish behavior as compared against both groups of lower moral identity centrality participants as well as higher centrality participants who were assigned to the control group.

We analyzed whether the patterns shown in Figure 2 followed the predicted linear trends of (a) decreasing cooperation relative to the initial trials among those who were morally primed, but whose moral identity was not highly central and (b) sustained cooperation among those who are morally primed and whose moral identity has high centrality. We tested these predictions by first performing a median split on the moral identity measure to create four groups consisting of primed higher centrality participants ( $n = 8$ ), primed lower centrality participants ( $n = 7$ ), nonprimed higher centrality participants ( $n = 8$ ), and nonprimed lower centrality participants ( $n = 10$ ). We then conducted a repeated measures ANOVA on cooperation across the five trial blocks within each of these four groups. We tested the significance of a linear contrast within each group to assess whether there was a change in cooperation over time. Our results showed that the linear contrast was significant in lower centrality/primed,  $F(1, 6) = 6.99, p < .05$ , and higher centrality/nonprimed groups,  $F(1, 7) = 16.75, p < .05$ , but not in the other two groups. We interpret this result as supporting our predictions because the results showed that cooperation declines significantly with time among unprimed higher moral identity centrality participants and primed lower centrality participants but remains unchanged over time in the other two groups. That is, cooperation remains consistently high in the high-moral identity/primed group and consistently low in the low-moral identity/unprimed group.

### Discussion

Study 4 integrated and extended the findings from our previous studies by examining the influence of two types of situational cues: a moral prime that should motivate prosocial behavior and feedback about the selfish behavior of others that should motivate self-interested actions. Results showed that participants for whom moral identity has higher centrality initially exhibited higher levels of cooperative behavior than lower centrality participants, but feedback about the selfish behavior of others quickly reduced their level of cooperation. Among lower centrality participants, results showed that the presence of a moral prime initially increased cooperative behavior, but once again, this effect lasted for only one trial block. Only the combination of higher as opposed to lower centrality of moral identity and the presence of a moral prime resulted in sustained cooperative behavior in the face of the selfish actions of others. This pattern of findings is consistent with the notion that the influence of moral identity on behavior is contingent on the current accessibility of moral identity within the working self-concept relative to other potentially competing aspects of identity.

### General Discussion

In their review of the extant literature on moral identity, Hardy and Carlo (2005) noted that the concept holds great promise for

advancing researchers' present understanding of moral functioning, but many fundamental questions remain unanswered. For example, relatively little is known about the mechanisms through which moral identity influences moral action, and there is a paucity of empirical studies examining potential mediating and moderating factors (Hardy & Carlo, 2005). More recently, Walker and Frimer's (2007) investigation of different types of moral exemplars (i.e., caring vs. brave) highlighted the need for further research aimed at determining how situational and personal variables interact to influence behavior in the moral domain. Through the adoption of a social-cognitive perspective on the centrality of moral identity and the empirical assessment of hypotheses derived from the resulting theoretical framework, this article has advanced present knowledge regarding how and when situational factors interact with moral identity centrality to jointly influence moral outcomes.

One important contribution of our research relates to its specification and testing of a mediating mechanism—the current accessibility of moral identity within the working self-concept—through which situational factors and moral identity centrality influence moral intentions and behaviors. According to social-cognitive principles, the accessibility of a knowledge structure, such as the moral self-schema, determines its potential to be used in processing and acting on information (cf. Higgins & Brendl, 1995). Thus, the accessibility of moral identity within the working self-concept should determine the extent to which it influences moral outcomes. This hypothesis was supported in Study 1, which showed that as the current accessibility of moral identity increased, intentions to behave in a prosocial manner also increased. It was also supported in Study 2, which showed that as the current accessibility of moral identity decreased, intentions to behave in a selfish manner increased.

Establishing the current accessibility of moral identity as a proximal determinant of moral outcomes also allowed us to theorize about how and when situational factors can be expected to influence moral actions. We hypothesized that situational cues like moral primes that activate the moral self-schema can promote prosocial intentions and behavior by increasing the current accessibility of moral identity. Study 1, which included a measure of the current accessibility of moral identity, provided direct support for this prediction; Study 4 findings provided inferential support for current accessibility as a mediating mechanism. We also hypothesized that situational cues like financial incentives for task performance and feedback about the selfish behavior of others would decrease prosocial intentions and behavior by decreasing the current accessibility of moral identity. Study 2 provided direct support for this prediction; Studies 3 and 4 provided inferential support. The notion that situational factors can make certain identities more or less accessible, and thereby influence behavior, has been discussed at a theoretical level (cf. Aquino et al., 2005; Skitka, 2003). However, this assertion has been tested in few empirical studies. Thus, our assessment of hypothesized links between specific situational factors and moral outcomes contributes to present understandings of moral functioning.

A third contribution of our research relates to the incorporation of models of human goals (Grouzet et al., 2005) and values (Schwartz, 1992, 1994; Schwartz & Boehnke, 2004) into a social-cognitive framework. Drawing from the notion that some goals/values are inherently oppositional to others (e.g., self-enhancement



values conflict with self-transcendent values; see Burroughs & Rindfleisch, 2002; Kasser et al., 2007), we hypothesized that the presence of a self-interest-promoting situational cue would create a dissonant psychological state for people for whom moral identity is relatively high in centrality that would be resolved through a reduction in the current accessibility of moral identity. Consistent with this prediction, Study 2 showed that a financial incentive for task performance increased the current accessibility of achievement-oriented facets of identity and also decreased the current accessibility of moral identity. Notably, the decrease in the current accessibility of moral identity was most pronounced for participants for whom moral identity was relatively high in centrality. Although Studies 3 and 4 did not include measures of current accessibility, study findings were also consistent with the notion that self-interest-promoting situational factors decreased the current accessibility of moral identity. This was especially true for participants for whom moral identity was higher in centrality, as self-interested, financial cues were shown to have a stronger influence on their behaviors compared with their lower moral identity centrality counterparts.

At first glance, the observation that situational cues can promote self-interested behavior even among people for whom moral identity has high centrality (Studies 3 and 4) appears to contradict a large body of research that conceptualizes moral identity as enduring and stable (cf. Blasi, 1984, 2005; Colby & Damon, 1992, 1993; Damon, 1984; Damon & Hart, 1992; Walker & Frimer, 2007; Walker & Hennig, 2004). We believe it is possible to reconcile the social-cognitive view of the self-concept as dynamic and multifaceted with the notion that moral identity can be an enduring, a stable, and a deeply held aspect of the self-concept for some people. The basis for this reconciliation pertains to the regularity with which people for whom moral identity is highly central encounter moral primes and self-interest-promoting situational cues. For example, consider a person who volunteers at a nonprofit organization and regularly attends religious services. Such a person is likely to have his or her highly central moral self-concept continually activated and reinforced, thereby maintaining the accessibility of moral identity within the working self-concept and promoting sustained commitment to moral action. In contrast, consider a person who works in a highly competitive industry and faces constant pressure to deliver financial results. Such a person is likely to have self-achievement-oriented facets of identity continually activated, thereby reducing the accessibility of moral identity within the working self-concept, and inducing situational variability in his or her commitment to moral action. Although our studies do not provide a rigorous test of this cumulative experiences-based explanation for predicting enduring versus variable commitment to moral action, findings from Study 4 are consistent with this rationale. Without a moral prime, participants for whom moral identity was highly central did not sustain an increased level of cooperative behavior in the face of selfish behavior by others. However, with a moral prime, such participants did sustain cooperative behavior over time. Arguably, these findings point to the importance of having salient situational cues available to reinforce moral identity when other situational cues might otherwise deactivate it within the working self-concept.

The studies reported herein are subject to three notable limitations, each of which represents an important avenue for further research. First, only a small number of specific situational factors

were examined in our studies. The full range of situational factors that may activate the moral self-schema or competing facets of identity is unknown and multifarious. For those who may be interested in promoting moral action, the limited number of situations we studied is particularly lamentable because neither of the moral primes used in our studies (i.e., recalling and reading a list of the Ten Commandments and completing a handwriting task) would seem to have practical utility. One specific type of moral prime that may prove useful in fostering moral behavior involves witnessing the morally exemplary actions of others (cf. Freeman, Aquino, & McFerran, 2009). Because exemplary actions may be witnessed in social settings or business contexts (Shao et al., 2008), the examination of this type of moral prime may substantially advance present understandings of everyday moral functioning.

A second limitation of this research relates to our hypothesis regarding the influence of self-interest-promoting situational factors on people for whom moral identity has relatively high centrality. This hypothesis was based on the notion that such factors create a dissonant psychological state wherein moral identity and self-enhancement-related facets of identity are simultaneously active within the working self-concept. Consequently, the accessibility of moral identity is predicted to decrease as a means of dissonance reduction. Although there is ample evidence to support our rationale (cf. Burroughs & Rindfleisch, 2002; Kasser et al., 2007), the direct assessment of aroused psychological tensions resulting from self-interest-promoting situational factors would further increase confidence in the veracity of this explanation.

A third limitation relates to the need for self-consistency. As Blasi (1980, 1984) pointed out, people for whom moral identity is highly central should experience discomfort when they violate their own moral standards. One interesting question that our studies do not address is what happens when people with highly central moral identities behave in a selfish fashion (as they were induced to do by self-interest-promoting situational cues in Studies 3 and 4). It may be that such people experience greater distress than those for whom moral identity has lower centrality. Alternatively, people with highly central moral identities may execute elaborate forms of cognitive rationalization to avoid self-condemnation. Future research should address this question to better understand the psychological and emotional consequences of violating moral standards given our results showing that few of us are either unflinching saints or irredeemable sinners, but rather something in between.

## References

- Aiken, L. S., & West, S. G. (1991). *Multiple regression: Testing and interpreting interactions*. Thousand Oaks, CA: Sage.
- Aquino, K. (1998). The effects of ethical climate and the availability of alternatives on the use of deception during negotiation. *International Journal of Conflict Management*, 9, 195–217.
- Aquino, K., & Freeman, D. (in press). Moral identity in business situations: A social-cognitive framework for understanding moral functioning. In D. Narvaez & D. Lapsley (Eds.), *Personality, identity, and character: Explorations in moral psychology*. New York: Cambridge University Press.
- Aquino, K., & Reed, A., II. (2002). The self-importance of moral identity. *Journal of Personality and Social Psychology*, 83, 1423–1440.
- Aquino, K., Reed, A. II, Stewart, M., & Shapiro, D. (2005). Self-regulatory

- identity theory and reactions toward fairness enhancing organizational policies. In S. W. Gilliland, D. D. Steiner, D. P. Skarlicki, & K. Van den Bos (Eds.), *What motivates fairness in organizations? Research in social issues in management* (pp. 129–148). Charlotte, NC: Information Age Publishing.
- Aquino, K., Reed, A., II, Thau, S., & Freeman, D. (2007). A grotesque and dark beauty: How the self-importance of moral identity and mechanisms of moral disengagement influence cognitive and emotional reactions to war. *Journal of Experimental Psychology*, *43*, 385–392.
- Bandura, A. (2001). Social cognitive theory: An agentic perspective. *Annual Review of Psychology*, *52*, 1–26.
- Batson, C. D., Kobryniewicz, D., Dinnerstein, J. L., Kampf, H. C., & Wilson, D. W. (1997). In a very different voice: Unmasking moral hypocrisy. *Journal of Personality and Social Psychology*, *72*, 1335–1348.
- Bem, S. L. (1981). Gender schema theory: A cognitive account of sex typing. *Psychological Review*, *88*, 354–364.
- Blasi, A. (1980). Bridging moral cognition and moral action: A critical review of the literature. *Psychological Bulletin*, *88*, 1–45.
- Blasi, A. (1984). Moral identity: Its role in moral functioning. In W. Kurtines & J. Gewirtz (Eds.), *Morality, moral behavior and moral development* (pp. 128–139). New York: Wiley.
- Blasi, A. (1993). The development of identity: Some implications for moral functioning. In G. G. Noam, T. E. Wren, G. Nunner-Winkler, & W. Edelstein (Eds.), *The moral self* (pp. 99–122). Cambridge, MA: MIT Press.
- Blasi, A. (2004). Moral functioning: Moral understanding and personality. In D. K. Lapsley & D. Narvaez (Eds.), *Moral development, self, and identity* (pp. 335–348). Mahwah, NJ: Erlbaum.
- Blasi, A. (2005). Moral character: A psychological approach. In D. K. Lapsley & F. C. Power (Eds.), *Character psychology and character education* (pp. 67–100). Notre Dame, IN: University of Notre Dame Press.
- Bok, S. (1978). *Lying: Moral choice in public and private life*. New York: Vintage Books.
- Burke, P. J. (2003). Relationships among multiple identities. In T. J. Owens, R. T. Serpe, & P. A. Thotis (Eds.), *Advances in identity theory and research* (pp. 195–214). New York: Kluwer Academic/Plenum Publishers.
- Burroughs, J. E., & Rindfleisch, A. (2002). Materialism and well-being: A conflicting values perspective. *Journal of Consumer Research*, *29*, 348–370.
- Bussey, K. (1999). Children's categorization and evaluation of different types of lies and truths. *Child Development*, *70*, 1338–1347.
- Carver, C. S., & Scheier, M. F. (1998). *On the self-regulation of behavior*. New York: Cambridge University Press.
- Cervone, D., & Shoda, Y. (1999). *The coherence of personality: Social-cognitive bases of consistency, variability, and organization*. New York: Guilford Press.
- Colby, A., & Damon, W. (1992). *Some do care: Contemporary lives of moral commitment*. New York: Free Press.
- Colby, A., & Damon, W. (1993). The uniting of self and morality in the development of extraordinary moral commitment. In G. Noam & T. Wren (Eds.), *The moral self* (pp. 149–174). Cambridge, MA: MIT Press.
- Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review*, *82*, 407–428.
- Conway, M. A., & Pleydell-Pearce, K. (2000). The construction of autobiographical memories in the self-memory system. *Psychological Review*, *107*, 261–288.
- Damon, W. (1984). Self-understanding and moral development from childhood to adolescence. In W. Kurtines & J. Gewirtz (Eds.), *Morality, moral behavior and moral development* (pp. 109–127). New York: Wiley.
- Damon, W., & Hart, D. (1992). Self-understanding and its role in social and moral development. In M. Bornstein & M. E. Lamb (Eds.), *Developmental psychology: An advanced textbook* (3rd ed., pp. 421–464). Hillsdale, NJ: Erlbaum.
- 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*, 100–119.
- Deshpande, R., Hoyer, W. D., & Donthu, N. (1986). The intensity of ethnic affiliation: A study of the sociology of hispanic consumption. *Journal of Consumer Research*, *13*, 214–220.
- 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*, *92*, 374–391.
- Doris, J. M. (1998). *Persons, situations, and virtue ethics*. London: Blackwell Publishers.
- Eisenberg, N. (2000). Emotion, regulation, and moral development. *Annual Review of Psychology*, *51*, 665–697.
- Freeman, D., & Aquino, K. (2008). *Evidence supporting the use of the moral identity internalization scale as a valid measure of moral identity centrality*. Unpublished manuscript.
- Freeman, D., Aquino, K., & McFerran, B. (2009). Overcoming beneficiary race as an impediment to charitable donations: Social dominance orientation, the experience of moral elevation, and donation behavior. *Personality and Social Psychology Bulletin*, *35*, 72–84.
- Gilligan, C. (1982). *In a different voice*. Cambridge, MA: Harvard University Press.
- Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology*, *74*, 1464–1480.
- Grouzet, F. M. E., Kasser, T., Ahuvia, A., Dols, J. M. F., Kim, Y., Lau, S., et al. (2005). The structure of goal contents across 15 cultures. *Journal of Personality and Social Psychology*, *89*, 800–816.
- Hardy, S. A., & Carlo, G. (2005). Identity as a source of moral motivation. *Human Development*, *48*, 232–256.
- Harman, G. (2003). No character or personality. *Business Ethics Quarterly*, *13*, 87–94.
- Higgins, E. T. (1989). Knowledge accessibility and activation: Subjectivity and suffering from unconscious sources. In J. S. Uleman & J. A. Bargh (Eds.), *Unintended thought* (pp. 75–123). New York: Guilford Press.
- Higgins, E. T. (1996). The “self digest”: Self-knowledge serving self-regulatory functions. *Journal of Personality and Social Psychology*, *71*, 1062–1083.
- Higgins, E. T., & Brendl, C. M. (1995). Accessibility and applicability: Some “activation rules” influencing judgment. *Journal of Experimental Social Psychology*, *31*, 218–243.
- Kant, I. (1959). *Foundation of the metaphysics of morals* (L. W. Beck, Trans.). Indianapolis, IN: Bobbs-Merrill. (Original work published 1785)
- Kasser, T., Cohn, S., Kanner, A. D., & Ryan, R. M. (2007). Some costs of American corporate capitalism: A psychological exploration of value and goal conflicts. *Psychological Inquiry*, *18*, 1–22.
- Kay, A., Wheeler, S. C., Bargh, J. A., & Ross, L. (2004). Material priming: The influence of mundane physical objects on situational construal and competitive behavioral choice. *Organizational Behavior and Human Decision Processes*, *95*, 83–96.
- Kelley, H. H., Holmes, J. G., Kerr, N., Reis, H., Rusbult, C. E., & Van Lange, P. A. (2003). *An atlas of interpersonal situations*. Cambridge, England: Cambridge University Press.
- Kihlstrom, J. F., & Klein, S. B. (1994). The self as a knowledge structure. In R. S. Wyer Jr. & K. Thomas (Eds.), *Handbook of social cognition: Vol. 1. Basic processes* (2nd ed., pp. 153–208). Hillsdale, NJ: Erlbaum.
- Kohlberg, L. (1969). Stage and sequence: The cognitive developmental approach to socialization. In D. A. Goslin (Ed.), *Handbook of socialization theory and research* (pp. 347–480). Chicago, IL: Rand McNally.

- Komorita, S. S., & Parks, C. D. (1994). *Social dilemmas*. Madison, WI: Brown and Benchmark.
- Lapsley, D. K. (1996). *Moral psychology*. Boulder, CO: Westview Press.
- Lapsley, D. K., & Lasky, B. (2001). Prototypic moral character. *Identity, 1*, 345–363.
- Lapsley, D. K., & Narvaez, D. (2004). A social-cognitive approach to the moral personality. In D. K. Lapsley & D. Narvaez (Eds.), *Moral development, self, and identity* (pp. 189–212). Mahwah, NJ: Erlbaum.
- Liebrand, W. B. G., & McClintock, C. G. (1988). The ring measure of social values: Computerized procedure for assessing individual differences in information processing and social value orientation. *European Journal of Personality, 2*, 217–230.
- Markus, H. (1977). Working self-concept and processing information about the self. *Journal of Personality and Social Psychology, 35*, 63–78.
- Markus, H., & Kunda, Z. (1986). Stability and malleability of the self-concept. *Journal of Personality and Social Psychology, 51*, 858–866.
- Mazar, N., Amir, O., & Ariely, D. (2008). The dishonesty of honest people: A theory of self-concept maintenance. *Journal of Marketing Research, 45*, 633–644.
- Midlarsky, E., Fagin Jones, S., & Corley, R. (2005). Personality correlates of heroic rescue during the Holocaust. *Journal of Personality, 73*, 907–934.
- Minsky, M. (1988). *The society of mind*. New York: Simon & Schuster.
- Mischel, W. (1968). *Personality and assessment*. New York: Wiley.
- Muller, D., Judd, C. M., & Yzerbyt, V. Y. (2005). When moderation is mediated and mediation is moderated. *Journal of Personality and Social Psychology, 89*, 852–863.
- Narvaez, D., Lapsley, D. K., Hagele, S., & Lasky, B. (2006). Moral chronicity and social information processing: Tests of a social cognitive approach to the moral personality. *Journal of Research in Personality, 40*, 966–985.
- Offerman, T., Sonnemans, J., & Schram, A. (1996). Value orientations, expectations and voluntary contributions in public goods. *The Economic Journal, 106*, 817–845.
- Peterson, C. C., Peterson, J. L., & Seeto, D. (1983). Developmental changes in ideas about lying. *Child Development, 54*, 1529–1535.
- Pillutla, M. M., & Chen, X. P. (1999). Social norms and cooperation in social dilemmas: The effects of context and feedback. *Organizational Behavior and Human Decision Processes, 78*, 81–103.
- Reed, A., II, & Aquino, K. (2003). Moral identity and the expanding circle of moral regard toward out-groups. *Journal of Personality and Social Psychology, 84*, 1270–1286.
- Reed, A., II, Aquino, K., & Levy, E. (2007). Moral identity and judgments of charitable behaviors. *Journal of Marketing, 71*, 178–193.
- 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*, 1610–1624.
- Ross, L., & Nisbett, R. E. (1991). *The person and the situation: Perspectives of social psychology*. (McGraw-Hill series in social psychology). New York: McGraw-Hill.
- Sage, L., Kavussanu, M., & Duda, J. (2006). Goal orientations and moral identity as predictors of prosocial and antisocial functioning in male association football players. *Journal of Sports Sciences, 24*, 455–466.
- Schwartz, S. H. (1992). Universals in the content and structure of values: Theory and empirical tests in 20 countries. In M. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 25, pp. 1–65). New York: Academic Press.
- Schwartz, S. H. (1994). Are there universal aspects in the content and structure of values? *Journal of Social Issues, 50*, 19–45.
- Schwartz, S. H., & Boehnke, K. (2004). Evaluating the structure of human values with confirmatory factor analysis. *Journal of Research in Personality, 38*, 230–255.
- Shao, R., Aquino, K., & Freeman, D. (2008). Beyond moral reasoning: A review of moral identity research and its implications for business ethics. *Business Ethics Quarterly, 18*, 513–540.
- Singer, P. (1981). *The expanding circle: Ethics and sociobiology*. New York: Farrar Straus & Giroux.
- Skitka, L. J. (2003). Of different minds: An accessible identity model of justice reasoning. *Personality and Social Psychology Review, 7*, 286–297.
- Vohs, K. D., Mead, N. L., & Goode, M. (2006, November 17). The psychological consequences of money. *Science, 314*, 1154–1156.
- Walker, L. J., & Frimer, J. A. (2007). Moral personality of brave and caring exemplars. *Journal of Personality and Social Psychology, 93*, 845–860.
- Walker, L. J., & Hennig, K. H. (2004). Differing conceptions of moral exemplarity: Just, brave, and caring. *Journal of Personality and Social Psychology, 86*, 629–647.
- Walker, L. J., & Pitts, R. C. (1998). Naturalistic conceptions of moral maturity. *Developmental Psychology, 34*, 403–419.
- Weaver, G. R. (2006). Virtue in organizations: Moral identity as a foundation for moral agency. *Organization Studies, 27*, 341–368.

Received March 14, 2006

Revision received January 7, 2009

Accepted January 13, 2009 ■