Journal of Experimental Psychology: General

# The End Justifies the Means, but Only in the Middle

Maferima Touré-Tillery and Ayelet Fishbach

Online First Publication, October 17, 2011. doi: 10.1037/a0025928

Journal of Experimental Psychology: General

0096-3445/11/$12.00 DOI: 10.1037/a0025928

# The End Justifies the Means, but Only in the Middle

Maferima Touré-Tillery and Ayelet Fishbach
University of Chicago

Achieving goals often requires the completion of sequential actions, such as finishing a series of assignments to pass a class. In the course of pursuing such goals, people can decide how closely to follow their personal standards for each action. We propose that actions at the beginning and end of a sequence appear more diagnostic of the pursuer's personal standards than do actions in the middle. Therefore, people are more likely to adhere to their standards at the beginning and end of goal pursuit—and slack in the middle. We demonstrate this pattern of judgment and behavior in adherence to ethical standards (e.g., cheating), religious traditions (e.g., skipping religious rituals), and performance standards (e.g., "cutting corners" on a task). We also show that the motivation to adhere to standards by using proper means is independent and follows a different pattern from the motivation to reach the end state of goal pursuit.

*Keywords:* motivation, goals, standards, ethics, self-signaling

People often pursue goals that require the completion of a sequence of actions. For example, a student must complete a series of assignments to pass a class, a cook needs to prepare several courses to make a meal, and a consumer with a frequent buyer card has to make a series of purchases before getting a reward. In the course of pursuing such goals, people can decide how closely to follow their personal standards for each action in the sequence. For example, the student can determine how closely to follow his or her ethical standards on each assignment, and the cook can decide how closely to follow his or her performance standards when preparing each course of the meal.

We define standards as norms or principles that people use to evaluate their own and others' conduct. In the context of goal pursuit, adherence to standards reflects a concern with using proper means rather than a focus on reaching the goal's end state. In general, adhering to standards sends positive signals to the self and others about a person's values and character. In some cases, people fail to adhere to their standards because they have not fully acquired the conventions of their society (Bandura, 1977; Burton & Kunce, 1995; Hoffman, 1977; Wright, 1971). Very often, however, individuals who are fully aware and capable of meeting these conventions succumb to the temptation to relax their standards to save time and effort or to gain some other valuable resource (Eisenberg & Shell, 1986).

Standards can apply to a variety of domains, ranging from ethics to task performance. The literature on ethics has identified several factors that can make ethical acts seem more acceptable and thereby lead to the relaxation of ethical standards. Thus, perceptions of wealth-based inequity (Gino & Pierce, 2009), unethical acts of ingroup members (Gino, Ayal, & Ariely, 2009), and depletion of self-control resources (Mead, Baumeister, Gino, Schweitzer, & Ariely, 2009) can all encourage people to relax their moral standards, at least temporarily. Moreover, increasing belief in determinism (as opposed to free will; Vohs & Schooler, 2008) or establishing one's moral credentials by expressing virtuous attitudes or traits (Monin & Miller, 2001; Sachdeva, Iliev, & Medin, 2009) can increase the likelihood that one will behave unethically.

The question we explore in this article is whether the position of an action in a sequence of goal-related actions influences the likelihood that individuals will adhere to (vs. relax) their standards when performing this action. Specifically, we explore the possibility that people follow their ethical and performance standards more carefully for actions at the beginning and end of a sequence because these actions provide a better signal to the self of the true nature of the self.

## The Costs of Relaxing Standards

Although relaxing standards carries obvious benefits, it can also carry costs associated with the consequences of being caught. Indeed, if caught relaxing certain types of standards, one can incur mild to severe costs ranging from social disgrace to financial losses. Rational crime theory suggests that people will relax their ethical standards (i.e., break laws) when a cost-benefit analysis reveals that the potential gains from relaxing standards outweigh such external costs (Allingham & Sandmo, 1972; Becker, 1968; Lewicki, 1984). This cost-benefit analysis perspective can also apply to other types of standards, such as adhering to performance standards or religious traditions. For example, relaxing performance standards on the job can have serious financial and social ramifications, and religious transgressions can often result in severe social consequences such as excommunication.

Beyond these external—social and financial—costs, relaxing one's standards can have internal, psychological costs even when

We are grateful to Tal Eyal, Wilhelm Hofmann, and Ann McGill for their help and thoughtful suggestions at various stages in the development of this work. We also thank Alex Simon and Sylvie Armand for their indispensable research assistance.

Correspondence concerning this article should be addressed to Maferima Touré-Tillery, Booth School of Business, University of Chicago, 5807 South Woodlawn Avenue, Chicago, IL 60637, or to Ayelet Fishbach, Booth School of Business, University of Chicago, 5807 South Woodlawn Avenue, Chicago, IL 60637. E-mail: mtt@chicagobooth.edu or ayelet.fishbach@chicagobooth.edu

no one is watching or judging. Private transgressions can have a negative impact on self-image because they can signal to one that one has low standards. Indeed, people learn about themselves in the same way they learn about others: by making inferences from observing their own actions. Such self-observations inform people of their own attitudes (Bem, 1972; Festinger, 1957) and goals (Fishbach, Dhar, & Zhang, 2006; Fishbach, Zhang, & Koo, 2009), and we assume these self-observations also inform people of their own standards. Moreover, inferences of self-standards can have important consequences for self-esteem because transgressions (e.g., cheating, slacking on the job) may suggest a failure to live up to an ideal level of conduct, thus threatening one's sense of self (Baumeister & Neuman, 1994; Eidelman & Biernat, 2007; Gollwitzer, Wicklund, & Hilton, 1982; Steele, 1988). For example, a student who cheats on an exam may interpret her behavior as a sign that she is dishonest, which could negatively affect her self-image because she—like most people—wishes to think of herself as an honest person.

Accordingly, people's search for evidence when making inferences about their self-standards tends to be biased toward supportive evidence, in a process captured by Baumeister and Neuman's (1994) metaphor of the "intuitive lawyer," whose search for evidence is motivated by a desired conclusion. To protect their self-image, people may selectively notice and remember the times they adhered to (vs. relaxed) their standards and, thus, gather information to confirm their desired self-conception of having high standards of ethics or performance (Swann & Read, 1981). Furthermore, people can maintain the perception of a moral self by avoiding comparisons between their questionable behaviors and known moral standards or by conveniently redefining those standards to make their behaviors seem more acceptable. For example, Batson, Thompson, Seuferling, Whitney, and Strongman (1999) found that participants were prone to deceive when reporting the outcome of a coin flip unless moral standards were salient enough to force behavior-standard comparisons. Mazar, Amir, and Ariely (2008) documented people's tendency to cheat "just a little bit"—enough to benefit from their dishonesty (e.g., getting more money from an experiment) but not enough to code their behavior as an ethical transgression that would hurt their self-image. Indeed, in these experiments, participants claimed much less money than the maximum amounts possible had they fully cheated. In addition, Zhong, Bohns, and Gino (2010) found an increase in unethical behavior in dimly (vs. brightly) lit settings where participants could presumably maintain an illusory sense of anonymity and perceive their actions as less noticeable.

These findings suggest people actively monitor their positive impressions of themselves by behaving (and thinking) in ways that allow them to signal to themselves that they have high standards. We adopt the term *self-signaling* to describe the process whereby an individual engages in an action at least in part for its diagnostic value, wishing to infer high self-standards from this action (see Greenwald & Breckler, 1985; Hogan, Jones, & Cheek, 1985; Prelec & Bodner, 2003; Schlenker, 1985).

Previous research suggests some actions are more diagnostic than others for inferences about the self and others; that is, they have higher signaling value. Specifically, actions that statistically deviate from the norm set by the majority and, hence, from expectations tend to be seen as more diagnostic of the pursuer's personal characteristics. For example, Skowronski and Carlston (1987) found that negative cues have a greater impact on interpersonal judgments than do positive and moderate ones because the former deviate from the norm, whereas the latter do not. The assumption is that most people under most circumstances conform to the norm, thus any deviation makes a statement about the individual (Skowronski & Carlston, 1989). For example, taking a daily jog in spite of the rain is more likely to be interpreted as a (positive) signal of willpower than is jogging on a sunny spring day because few people jog in bad weather. On the other hand, taking a private drink before noon may be seen as more indicative of an emerging alcoholism problem (a negative signal) than is drinking at dinnertime, since most people do not drink before noon (Bodner & Prelec, 1996). Self-signaling concerns might induce the jogger to go out on that rainy day, whereas the wine enthusiast might wait awhile before taking that drink.

Beyond these statistical considerations, we propose that the diagnostic value of actions fluctuates over the course of pursuing sequential actions toward a goal and that people are more motivated to meet their standards (i.e., self-signal high standards) when performing actions that have high diagnostic value. In particular, we explore the possibility that the position of an action in a sequence of goal-related actions (beginning vs. middle vs. end) influences the signaling value of such action, and hence influences adherence to standards.

## Relaxing Standards in the Middle

Previous research has demonstrated that beginning and end positions in a sequence are more distinctive than are middle ones, such that stimuli appearing at the beginning and end stand out from the rest. Murdock (1960) argued that this differential distinctiveness of stimuli in a sequence accounts for primacy and recency effects: the finding that when people are presented with a sequence of stimuli (e.g., words), they tend to remember the first few (primacy effect) and last few items (recency effect) better than those in the middle (see Greene, 1986, for review). We suggest the differential distinctiveness of sequential stimuli applies to actions in the course of goal pursuit, such that actions at the beginning and end are more salient than are actions in the middle.

Greater salience in turn leads to exaggerated judgments and perceptions of importance (Taylor & Fiske, 1978). For example, Taylor, Fiske, Etcoff, and Ruderman (1978) found that distinctive individuals (e.g., a solo female in a male group) receive a disproportionate amount of attention and are perceived as more influential to a group conversation than are nonsalient individuals. In a nonsocial context, Pryor and Kriss (1977) showed that a salient (vs. nonsalient) subject in a sentence was brought to mind more easily and was perceived as more responsible (causal) for the state of affairs that the sentence describes. In their experiments, the salience of the subject depended on its position relative to the object: salient subjects appeared before the object of the sentence (e.g., *John likes the car*), whereas nonsalient subjects came after the object (e.g., *The car is liked by John*). It is thus possible that in the context of self- and other-inferences, distinctiveness might also make beginning and end actions more diagnostic for inferences about the person pursuing the actions.

Because actions at the beginning and end of goal pursuit are salient, we hypothesize that people will—consciously or nonconsciously—perceive them as more diagnostic than actions in the

middle when making inferences about the traits and abilities of the person carrying out the action (self or other). Therefore, we predict that self-signaling concerns will drive people to adhere to their standards more closely at the beginning and end of a sequence of goal-related actions (where actions are more diagnostic) and to slacken their standards in the middle (where actions are less diagnostic). We expect these patterns of judgments and behaviors to occur in the pursuit of actions reflecting on ethical standards (e.g., not cheating), religious standards (e.g., following a religious ritual), and performance standards (e.g., producing high-quality outputs on a task).

We posit this hypothesis against a learning perspective, which would predict linear changes in adherence to standards over the course of goal pursuit as people gain experience with the contexts and tasks involved in pursuing the goal. In the performance domain, this perspective could mean an improving adherence to standards as people acquire skills at the beginning of a task and, hence, perform progressively better in the middle and at the end. This linear prediction for standard adherence could also apply to ethical standards, albeit in a different direction. People who are not fully aware of the implications of relaxing standards might start out adhering to their standards and then progressively relax them in the middle and at the end as they learn that breaking the rules has no real consequences—and as they get better at breaking them.

## Doing It Right Versus Getting It Done

We distinguish between two dimensions of motivation in the course of goal pursuit: the outcome-focused motivation to reach the end state ("getting it done") and the means-focused motivation to adhere to one's standards ("doing it right") in the process of reaching that end state. Most treatments of motivation patterns have focused on the motivation to get things done, or outcome-focused motivation, and have adopted the "psychophysics" perspective that the perceived marginal impact of goal actions drives fluctuations in this aspect of motivation. In particular, research on the goal gradient hypothesis suggests that the perceived marginal impact of actions increases with each consecutive action (Heath, Larrick, & Wu, 1999; Higgins & Brendl, 1995). For example, the last action accomplishes 100% of the remaining progress, which is twice the impact of the second-to-last action (i.e., 50%). As a result, the motivation to complete the goal increases monotonically with proximity to the goal's end state, such that people (and other animals) exert more effort and persistence as they approach the end state (Brown, 1948; Förster, Higgins, & Idson, 1998; Hull, 1932; Kivetz, Urminsky, & Zheng, 2006).

In contrast, our research examines the means-focused motivation to do things right, which is distinct from the outcome-focused motivation to get things done (see Higgins, Idson, Freitas, Spiegel, & Molden, 2003; Touré-Tillery & Fishbach, 2011). The psychophysical perspective (i.e., based on the perceived impact of actions) makes no predictions about the pattern of means-focused motivation because the extent to which people adhere to their standards does not influence the actual or perceived impact of their actions on accomplishing a goal. For example, whether a person cheats or not on a specific trial in a 10-trial task, this trial would still accomplish 10% of the task. In order to clearly distinguish between the motivation to reach the end state and our theorizing on the motivation to do it right, we purposely use goal configurations in which we expect the former to increase over time, following a goal-gradient pattern. We predict that within the same self-regulatory task, measures of outcome-focused motivation will follow a monotonically increasing pattern distinct from the u-shaped pattern of means-focused motivation.

We note that Bonezzi, Brendl, and De Angelis (2011) recently offered a revision to the goal gradient hypothesis, suggesting that the marginal impact of actions depends on whether individuals monitor their progress in terms of distance from the end state or from the initial state (see also Fishbach, Henderson, & Koo, 2011; Koo & Fishbach, 2008). When individuals focus on distance from the initial point, the perceived marginal impact of actions is also higher at the beginning of goal pursuit. For example, the first action accomplishes 100% of the progress to date, whereas the second action accomplishes only 50% of the progress. It follows that under certain conditions (i.e., when individuals switch their focus from initial to end state as they progress toward a goal), the psychophysics perspective predicts a u-shaped function for outcome-focused motivation. In these cases, it is possible that the greater motivation to get things done at the beginning and end states might come at the expense of the desire to do them right. Therefore, people might be more likely to relax their standards at the beginning and end (vs. middle) to ensure fast progress at those points. Against this alternative, we predict that the motivation to do things right will be greater at the beginning and end (vs. middle) of a sequence of goal-related actions, following fluctuations in the perceived diagnosticity of actions for self-inferences

## Present Research

We report five experiments that test whether position in a sequence of goal-related actions influences the degree of adherence to personal standards. We predict that people will adhere to their standards more closely at the beginning and end of goal pursuit and slack in the middle because they see actions at the beginning and end (vs. middle) as more diagnostic for inferences about the self and others. We tested our predictions in the contexts of ethical, religious, and performance standards and in situations in which adherence to standards constituted a secondary goal (i.e., behaving ethically while performing a series of tasks) or a focal goal (i.e., adhering to religious traditions during a religious holiday). Each experiment consisted of a series of actions toward a goal with clear beginning and end states.

In Experiments 1 and 2, we examined our hypothesis for ethical standards. Participants had the opportunity to cheat (i.e., misreport the outcome of a coin flip, take advantage of a forgetful experimenter) at different points along the sequence of actions. In Experiment 3, we explored our hypothesis for religious standards by measuring Jewish participants' adherence to the religious tradition of lighting Menorah candles during Hanukah, an 8-night Jewish holiday. Using religious standards, we could also explore the moderating role of commitment to a standard. In Experiments 4 and 5, we tested our hypothesis for performance standards. In addition, Experiment 4 explored the mediating role of the perceived diagnosticity of actions. Experiment 5 tested whether the motivation to reach a goal's end-state follows a pattern independent of that of the motivation to adhere to standards in the process of reaching that end state.

In all these experiments, except in the context of religious traditions (Experiment 3), we ensured complete privacy and anonymity, such that relaxing or adhering to standards had no external consequences in terms of self-presentation. By reducing social-signaling concerns, we sought to ensure that self-signaling considerations were the main drivers of participants' adherence to standards.

## Experiment 1: Flipping Coins

In Experiment 1, we examined whether position in a sequence influences adherence to ethical standards. Participants completed a series of 10 proofreading tasks. For each task, participants had to assign themselves to the short or long version of a passage by flipping a coin, presumably to ensure random allocation. We could infer dishonest behavior if the percentage of participants assigning themselves to the short proofreading task was greater than chance (50%, see Batson, Kobrynowicz, Dinnerstein, Kampf, & Wilson, 1997; Batson et al., 1999). We predicted that participants would be more likely to assign themselves to the short proofreading task in the middle of the experiment than at the beginning and at the end.

### Method

**Participants.** One hundred University of Chicago students (51 men, 49 women) participated in the experiment for monetary compensation. We excluded six participants from the analysis because they reported 9 or 10 favorable coin flips out of 10. The probabilities of getting such outcomes are smaller than 1% and .1%, respectively; thus, we suspected these participants were not striving to act ethically. In what follows, we report the analysis of data from the remaining 94 participants.

**Procedure.** This experiment employed a 10 (position of passage in the sequence: 1–10) within-subjects design. Participants completed the computerized experimental task in a private room where no one could observe them. They read that the task assessed "reading comprehension, verbal skills, and attention to detail for different types of passages with various writing styles, contents, and lengths." The participants' task was to proofread 10 different passages covering a wide variety of topics and containing spelling, grammatical, and other types of errors. They further read that for each passage, they would proofread either the short version (containing two errors) or the long version (containing 10 errors) and that for the purpose of the experiment, long and short passages had to be assigned randomly. Therefore, before reading each passage, participants had to flip a coin to determine which version to read. For this purpose, we provided participants with a coin labeled "SHORT" on one side and "LONG' on the other.

Using this paradigm, we created incentives for claiming the coin landed on SHORT (favorable outcome): reading a shorter passage, finding fewer errors, and completing the experiment faster. We labeled the coin to limit ambiguity, confusion (honest mistakes), or self-deception that could occur when people first flip the coin and then try to determine the meaning of the outcome (e.g., whether "heads" means they should read the short or long passage; see also Batson et al., 1997; Batson et al., 1999).

Finally, to emphasize the goal sequence, the instructions informed participants that they would get 10 points for each passage completed, regardless of the length of the passage, and that their goal was to proofread all 10 passages carefully to get a total of 100 points. To make sure participants were aware of their progress, we numbered each passage (e.g., "Passage 3 of 10") and displayed a progress chart before each passage, showing participants their present position (from Passage 1 through Passage 10), along with the number of points accumulated. After participants completed the proofreading tasks, the experimenter thanked and debriefed them.

### Results and Discussion

The coin-flipping paradigm allowed us to observe a continuous pattern of behavior. We measured deception by the proportion of participants who read the short version of the passage in each trial. If this proportion was greater than chance (50%) for a particular passage in the sequence, we could conclude that some participants dishonestly assigned themselves to the short passage to save time and effort.

We expected participants to be more honest about the outcome of the coin flip for trials at the beginning and at the end than for those in the middle, such that the ratio of short to long passages read would be roughly 1 to 1 for beginning and end trials. We performed a binomial test to determine whether the proportion of participants reporting favorable coin flips (i.e., coin landed on SHORT) on each passage was greater than 50%. As Figure 1 shows, these proportions were not significantly different from 50% for Passages 1 (50%), 2 (51%), 3 (59%), 5 (59%), and 10 (59%; $ps > .12$), suggesting that participants were honest in their reports of the coin flip outcome at the beginning and end of the task. However, as expected, the proportion of participants assigning themselves to the short passage was higher than chance for Passages 4 (61%), 6 (72%), 7 (68%), 8 (62%), and 9 (63%; $ps < .05$). With the exception of Passage 5, these results suggested that more deception had occurred in the middle (vs. beginning and end) of the task and provided initial evidence in support of our hypothesis.

To rule out the possibility that the observed pattern of aggregated data was a combination of two different linear patterns—that is, some participants cheated more at the beginning and others cheated more at the end—we conducted a within-subject analysis of the data, using a linear model for binary responses (Wooldridge, 2002). This analysis revealed a marginal effect of position on adherence to ethical standards, $F(8.5, 790.79) = 1.79$, $p = .07$, and the predicted quadratic contrast,

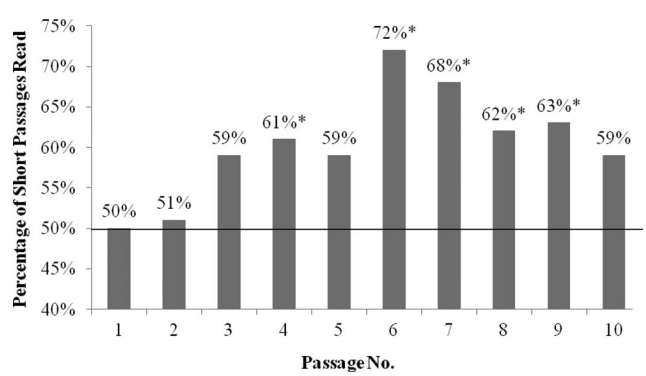


*Figure 1.* Percentages of participants who reported the favorable outcome of the coin flip and assigned themselves to the short proofreading passage, for each of the 10 passages in the sequence (Experiment 1). Horizontal line at 50% value indicates chance level. $^{*}p < .05$ (percentage greater than chance).

$F(1, 93) = 7.67$, $p < .01$. We conclude that people are more likely to relax their ethical standards in the middle (vs. beginning and end) of goal pursuit. We hypothesize that this pattern of behavior reflects their attempt to secure a positive self-image because actions at the beginning and end (vs. middle) of goal pursuit are seen as more diagnostic of one's ethics.

We note that in this experiment, participants might have adhered more closely to their standards at the beginning and end (vs. middle) partially because they were more aware of the presence of the experimenter at those points in the course of the task. Indeed, at the beginning, the experimenter had left the room just before participants started reading the instructions, whereas at the end, participants anticipated interacting with the experimenter again (to tell him that they had completed the experiment). Therefore, although they performed the tasks privately and knew their responses were anonymous, participants may still have adjusted their behaviors partially due to this recent or upcoming social interaction. We controlled for this alternative in the next experiment, using a paradigm in which participants interacted with the experimenter throughout the experiment. In that paradigm, greater adherence to ethical standards at the beginning and end should not reflect real or imaginary interactions with the experimenter on these trials. Furthermore, the next experiment moved to a paradigm where each participant had only one opportunity to cheat, at the beginning, middle, or end of the sequence. This paradigm controlled for the possibility that dishonesty increased after people built up an honest reputation at the beginning or learned they could get away with cheating when moving from the beginning to the middle of the task. The new procedure also controlled for the possibility that dishonesty decreased as people started feeling guilty over their recent transgressions when moving from the middle to the end of the task.

## Experiment 2: Accepting Undeserved Credit

In Experiment 2, participants had the opportunity to take advantage of the experimenter's "forgetfulness" and get undeserved credit for a trial they did not complete. We predicted participants would be more likely to accept the undeserved credit when the experimenter "mistakenly" offered it for a trial in the middle of the sequence than for one at the beginning or end of the seven-trial sequence.

### Method

**Participants.** Fifty-one University of Chicago students (28 men, 23 women) participated in the experiment for monetary compensation.

**Procedure.** This experiment employed a three (position in the sequence: beginning vs. middle vs. end) between-subjects design. In individual rooms, participants completed an experiment, presumably on how young adults color. Their task was to color seven images following the color scheme provided for each image. For this purpose, they received a set of coloring pens and a "coloring progress card." The progress card featured seven numbered boxes representing the seven images to color and was similar in design to frequent-buyer cards consumers often encounter in the marketplace (e.g., in coffee shops and bakeries).

The experimenter brought the images to participants one at a time. When a participant finished coloring an image, he or she would call the experimenter, who would stamp the progress card in the appropriate box (e.g., Box 1 for Image 1) and hand out the next image. Participants' goal was to get seven stamps on their progress cards to complete the experiment. To emphasize progress in the task, we numbered each image clearly (e.g., "Image 4 of 6") and left the progress card in the room with the participant.

At a critical point during the experiment, the experimenter pretended to forget he had already stamped the progress card and offered to stamp it again, thus providing a tempting opportunity for participants to do less work for the same compensation. Depending on the experimental condition, this "mistake" happened after Images 2, 4, or 6 (beginning, middle, or end, respectively). After a participant completed the critical trial, the experimenter stamped the card and left the room to get the next image. When he came back into the room, the experimenter would act confused, as if he thought that he had forgotten to stamp the card before leaving the room. The experimenter would look at the progress card and say, "Oh, it looks like I forgot to stamp your card before I left. . . . Can you give me your card so that I can mark it?" Then the experimenter would stamp the card again unless the participant stopped him and said that he or she had already received a stamp for that image.

Importantly, to encourage participants to correct the experimenter's mistake, on the first coloring trial of all experimental conditions, the experimenter would always "forget" to stamp the progress card. When the experimenter would come back into the room with the second image, he would announce, "Oh, I forgot to mark your progress card. I am supposed to do that before I bring you the next image." The experimenter would then mark the card before leaving the room. We used this procedure in all conditions to make participants expect mistakes and feel more comfortable about correcting the experimenter's critical "mistake" when he would offer the undeserved stamp later in the experiment. Furthermore, the fact that the experimenter appeared distracted and perhaps even incompetent should also have reduced self-presentation concerns in this particular context (Leary & Kowalski, 1990). Therefore, we could infer that participants who chose to accept the underserved stamps were not just respectful but rather deceitful (i.e., took advantage of the experimenter's "forgetfulness" to get credit for a task they did not complete). Using this procedure, the opportunity to cheat in the beginning condition came only after the second trial, which gave participants a chance to experience the task before the ethical dilemma presented itself. Upon completing the experiment, participants received appropriate debriefing information.

### Results and Discussion

Acceptance of undeserved stamps (dishonest behavior) followed the predicted pattern. Participants were less likely to correct the experimenter's mistake when it occurred in the middle than when it occurred at the beginning and end of the sequence. One-sided Fisher's exact tests confirmed that a greater percentage of participants accepted the extra stamp when the experimenter offered it after Image 4 (53%) than when the experimenter offered it after Images 2 (18%; $p < .05$) and 6 (18%; $p < .05$). Participants were more likely to relax their ethical standards and take advantage of

the experimenter's "forgetfulness" in the middle of the experiment than at the beginning and at the end. Because participants only had one opportunity to behave unethically, this pattern could not reflect ignorance of the nature of the task at the beginning, guilt about previous transgressions at the end, or random fluctuations in adherence to standards over the course of the task.

In our first two experiments, we found consistent evidence that people relax their ethical standards in the middle (vs. beginning and end) of goal pursuit, and we argue that this pattern of behavior occurs because people perceive beginning and end actions to be more diagnostic of their ethicality than middle actions. In our next experiment, we sought to provide evidence for this underlying mechanism by examining the moderating role of individual differences in levels of commitment to a given standard (i.e., the extent to which people hold and endorse a particular standard). We expected commitment to a particular standard to increase concern with self-signaling in the relevant domain; therefore, the effect of position on adherence to standard should increase with the level of commitment to the standard.

## Experiment 3: A Field Experiment: Adhering to Religious Traditions

Religious traditions are similar to ethical standards in the sense that they lead to inferences about a person's moral character, but unlike ethical standards, religious traditions are less universal. For example, whereas most people believe in the importance of being honest, not everyone values adherence to specific religious rules (e.g., dietary or clothing restrictions). Thus we predicted that participants' levels of religiousness would moderate standard adherence over the course of pursuing a religious tradition: the more religious participants are, the more likely they will be to adhere to their religious standards at the beginning and end of a sequence compared with the middle.

Experiment 3 explored this hypothesis in a naturalistic context. The Jewish holiday, Hanukah, lent itself very well to our investigation, as it requires the lighting of candles (Menorah) every night for 8 consecutive nights. We sought to demonstrate that more people would adhere to the ritual on the first and last (vs. middle) nights and that people's degree of religiousness—which reflects the self-relevance of the standard—moderates this u-shaped pattern of standard adherence.

In the first (and main) part of the experiment (3A), we examined the effect of position in the sequence of 8 nights on adherence to the religious ritual as a function of a person's religiousness. In the second part (3B), we tested whether people would also judge others as less religious for not adhering to this religious ritual at the beginning and end (vs. middle) of the holiday. Moving from self-inferences to inferences about others, we predicted that people would treat another person's beginning and end actions as a better signal of his or her character.

### Experiment 3A: Following Religious Traditions

Participants in this experiment indicated whether they lit the Menorah candles on each of the 8 nights of Hanukah and rated their own degree of religiousness.

#### Method

**Participants.** Two hundred and two students (42 men, 160 women) at Ben Gurion University, Israel, participated in the experiment in return for candy prizes. All participants were Jewish, and all indicated that they observe Hanukah. Our sample did not include orthodox Jews, whom we assumed would observe the ritual on every single night.

**Procedure.** The experiment followed an eight (position of night in the sequence: 1–8) within-subjects design. We conducted a survey 2 days after Hanukah to minimize the chances participants would forget what they did during the 8-night holiday. Participants received a survey asking them to indicate whether they lit the Menorah on each night (1–8) of Hanukah. They could answer "yes," "no," or "don't remember."

We assessed participants' level of religiousness by asking them to rate the extent to which they keep kosher (i.e., adhere to the rules of eating only specific foods and handling them properly: 1 = *never*; 7 = *always*). We thought the extent of keeping kosher would be a good measure of religiousness because it is relatively unambiguous, as it involves specific dietary behaviors of which all of our participants were aware. In addition, because there are no particular beginning and end states to the pursuit of dietary restrictions, this behavior more likely reflects an ongoing adherence to religious traditions beyond the sequence of the Menorah ritual.

#### Results and Discussion

The pattern of adherence to religious standards confirmed our hypothesis. The position of a given night in the 8-night sequence predicted participants' Menorah-lighting behavior, that is, the likelihood that they would answer "yes"—as opposed to "no" or "don't remember"—to the question "Did you light the Menorah on night *X*?," $\chi^2(7, N = 1616) = 64.00, p < .001$. When we analyzed the proportion of participants who reported lighting the Menorah on each night (i.e., answered "yes" vs. "no" or "don't remember"), we found participants were more likely to engage in the Menorah-lighting ritual at the beginning and end than in the middle of the sequence. We summarize these results in Table 1. A within-subject

Table 1
*Percentages of Participants Who Reported Lighting the Menorah on Each Night of Hanukah (Experiment 3A)*

| | Sequence of 8 nights | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Night of Hanukah | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| Adherence % | $76_{a}$ | $55_{b,e}$ | $50_{b,c,e}$ | $48_{b,c}$ | $49_{b,d}$ | $43_{c,d}$ | $45_{c,d}$ | $57_{e}$ |

*Note*. Identical letters displayed in the subscript of the percentages indicate percentages that are not significantly different from each other.

analysis, with a linear model for binary responses (Wooldridge, 2002) confirmed the effect of position on adherence to religious standards, $F(7, 1407) = 12.47$, $p < .001$, yielding the predicted quadratic contrast, $F(1, 201) = 44.39$, $p < .001$, as well as a linear declining contrast, $F(1, 201) = 26.63$, $p < .001$.

These patterns of results further held when we analyzed the proportion of participants who answered "no," as opposed to "yes" or "don't remember," $\chi^2(7, N = 1616) = 38.7$, $p < .001$, suggesting that our findings cannot be explained in terms of better memory for action versus inaction (i.e., lighting vs. not lighting the candle). More people remembered lighting the Menorah at the beginning and end (vs. middle), and they also remembered specifically not engaging in the ritual in the middle (vs. beginning and end).

## Moderation Analysis

We conducted a logistic multilevel regression to examine the effect of religiousness on the pattern of adherence to standards across the sequence of 8 nights. We included the following independent variables: (a) position in the sequence of 8 nights as linear predictor, (b) position squared, to estimate the curvilinear effect, (c) religiousness (i.e., eating kosher), (d) the interaction of religiousness and position, and (e) the interaction of religiousness and position squared. The dependent variable was coded such that 0 indicated no adherence and 1 indicated adherence to the religious ritual.

Confirming the previous analyses, we found a quadratic main effect of position on adherence to standards ($\beta = .084$), $t(1610) = 6.87$, $p < .001$, combined with a linear main effect of position on adherence to standards ($\beta = -.71$), $t(1610) = -7.79$, $p < .001$. Thus, adherence to standards followed a u-shaped pattern over the course of Hanukah yet decreased somewhat over time. We also found a main effect of religiousness ($\beta = .26$), $t(200) = 3.34$, $p < .001$, such that, not surprisingly, the likelihood of engaging in the Menorah-lighting ritual was positively related to religiousness. Finally, as we expected, the analysis revealed that religiousness marginally moderated the quadratic effect described above ($\beta = .0095$), $t(1610) = 1.62$, $p = .053$ (one sided), indicating the u-shaped pattern of standard adherence was more pronounced for more religious participants.

For illustrative purposes, we conducted a correlation analysis with a series of eight binary logistic regressions to examine the relationship between religiousness and Menorah-lighting behavior on each of the 8 nights of Hanukah. We found that in general, religiousness was positively correlated with Menorah-lighting behavior at the beginning and end but not in the middle of Hanukah. As Figure 2 illustrates, the correlations between religiousness and lighting the Menorah candles were generally higher and statistically significant for beginning and end Nights 1, 2, 7, and 8 ($\beta$s = .23, .20, .29, & .20, respectively; $p$s $< .05$). However, for middle Nights 3, 4, 5, and 6, the correlations between religiousness and standard adherence were lower ($\beta$s = .13, .13, .16, & .11, respectively) and did not reach significance, except for Night 5 ($p < .05$).

Taken together, these analyses reveal a moderating role of religiousness in the relationship between position in the sequence of nights and adherence to religious standards, which offers further evidence for our hypothesis that adherence to standards depends on the diagnostic value of actions. Indeed, religious participants should care more than less religions participants about maintaining a religious self-image, and they achieved this image by adhering more closely to their standards at the beginning and end of Hanukah—and relaxing their standards in the middle, when their behavior was almost undistinguishable from that of less religious participants.

We realize this particular religious context might lend itself to social-signaling interpretations for participants' behaviors. Other people in a person's social circle might know whether he or she participated in the Menorah ritual on any given night, and this social component of goal pursuit, rather than self-signaling alone, could affect the person's behavior. However, previous experiments did not share this particular limitation and yet exhibited similar u-shaped patterns of adherence to standards over the course of goal pursuit. We argue that this pattern of behavior reflects people's

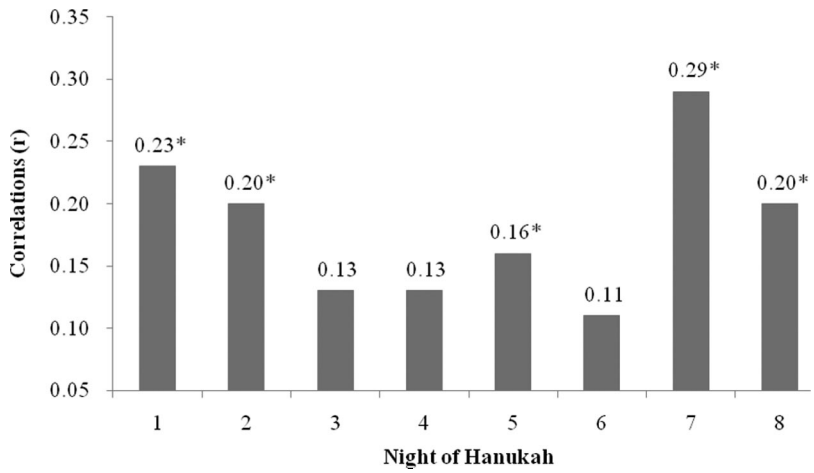


*Figure 2.* Correlations between religiousness and the lighting of the Menorah on each of the 8 nights of Hanukah (Experiment 3A). * $p < .05$ (statistically significant correlations).

attempt to secure a positive self-image at least to the same extent as it secures their public image.

## Experiment 3B: Impressions of Those Breaking Traditions

The Menorah ritual also provided a natural setting for us to investigate the diagnostic value of actions at different points along the sequence of nights, specifically, the differential inferences drawn about a person who fails to engage in the religious ritual at the beginning and end versus middle of Hanukah. We predicted that failing to engage in the Menorah ritual at the beginning and end (vs. middle) would be a stronger signal of a person's lack of religiousness. Furthermore, we assume the mechanisms for evaluating and learning about the self mirror those for judging and learning about others (Bem, 1972; Festinger, 1957). Therefore, finding that interpersonal judgments rely more heavily on others' first and last actions than on their middle actions would imply a similar pattern of judgment for the self.

Accordingly, in our follow-up experiment, we tested people's evaluations of another person's religiousness based on whether the latter failed to engage in the candle-lighting ritual on the first, middle, or last night of Hanukah.

### Method

**Participants.** Forty students (14 men, 26 women) from Ben Gurion University, Israel, participated in the experiment in return for candy prizes.

**Procedure.** The experiment employed a three (position of relaxing religious standards during Hanukah: beginning vs. middle vs. end) within-subjects design. All participants read three scenarios. In the first scenario, participants read about a person named Efrat who did not light the Menorah on the first night of Hanukah (beginning). To measure Efrat's perceived religiousness, we had participants rate how likely it is that Efrat "keeps tradition," "obeys religious laws," and "is religious" (1 = *very unlikely*; 10 = *very likely*). In the second and third scenarios, participants read about Shira and Hila, who did not light the Menorah on the fifth night (middle) and last night (end) of Hanukah, respectively. For each scenario, they answered the same questions about the actors' perceived religiousness. We provided no information about whether the actors lit the Menorah on the other seven nights.

### Results and Discussion

We averaged the three ratings of religiousness in each condition separately ($\alpha$s between .79 and .84). In support of our predictions, a one-way analysis of variance (ANOVA) revealed position in the sequence of nights influenced judgment of religiousness. Mauchly's test indicated the assumptions of sphericity had been violated, $\chi^2(2) = 15.07$, $p < .001$; therefore, we corrected degrees of freedom using Huynh-Feldt estimates of sphericity ($\varepsilon = .77$). The results showed position influenced judgment of religiousness, $F(1.55, 60.61) = 5.14$, $p < .01$.

Specifically, a quadratic contrast revealed that participants thought the persons who did not light the Menorah on the first and last night were less religious than the person who skipped the ritual on the fifth night, $F(1, 39) = 17.42$, $p < .001$. Further analysis revealed that participants perceived the person who did not light the Menorah on the first night ($M = 3.17$, $SD = 1.68$) as less religious than the one who failed to light the candle on the fifth night ($M = 3.63$, $SD = 1.78$), $t(39) = 3.19$, $p < .01$. In addition, participants perceived the person who did not engage in the ritual on the last night ($M = 3.35$, $SD = 1.82$) as less religious than the one who skipped it on the fifth night, $t(39) = 2.71$, $p < .01$. Finally, the difference in judgment of religiousness between the persons who did not light the Menorah on the first and last night was not significant, $t(39) = -1.02$, $p = .31$.

We conclude that people judge skipping religious traditions at the beginning and end of a goal sequence more harshly than they judge skipping traditions in the middle of that goal sequence. Because self-evaluation mirrors interpersonal evaluation, such a pattern of interpersonal judgment could imply that people weigh their own beginning and end actions more heavily than their middle actions. In the context of evaluating others, we further note that if people knew about the statistical distribution of standard adherence over the course of Hanukah, the inference that someone who skips the ritual on the first or last night is less religious than someone who skips it in the middle would be normative because such behavior is statistically rarer. However, it is unlikely that people are aware of our proposed u-shaped pattern of standard adherence.

Overall, the results of Experiment 3 support our hypothesis and generalize our findings to the context of religious standards. The experiment provides a field demonstration that individuals (especially those for whom the standard is self-relevant) are more likely to adhere to their religious standards at the beginning and end of goal pursuit than in the middle. In addition, we show that beginning and end actions are seen as more diagnostic in interpersonal judgments. Our next experiment served two distinct purposes. First, we sought to generalize our findings to performance standards and investigate whether people would be more likely to produce low-quality outputs in the middle of a task than at the beginning and end. Second, we intended to seek further evidence for our proposed underlying mechanism by exploring the mediating role of an action's perceived diagnosticity on the relationship between the position of the action and standard adherence. We chose to explore perceptions of diagnosticity in the context of performance because we suspected people would be more candid about reporting judgments of the self in this (somewhat benign) setting than in the context of ethics. Indeed, getting people to admit to cheating, let alone to make objective judgments about their own (dis)honesty, might be unrealistic.

## Experiment 4: Cutting Corners

In this experiment, we examined the influence of position in a sequence of goal-related actions on the diagnosticity of actions for self-inferences and on adherence to performance standards. Participants completed a moderately difficult shape-cutting task, and we coded the quality of their work (i.e., whether they literally cut corners). We expected participants to be more precise when cutting the first and last shapes than when cutting the shapes in the middle of the task.

To explore whether perceived diagnosticity of actions accounts for fluctuations in quality of work, upon completion of the cutting task, we provided participants with positive performance feedback on one of the shapes they had cut: first, middle, or last. We

predicted that participants who received positive feedback on their first or last shape would subsequently rate themselves as generally more skilled than those who received positive feedback on their third (middle) shape because beginning and end (vs. middle) actions are seen as more diagnostic of one's characteristics. In addition, we expected that differences in the perceived diagnosticity of beginning and end (vs. middle) actions would mediate the impact of position on performance.

## Method

**Participants.** Sixty students from various universities in the greater Chicago area (26 men, 34 women) participated in the experiment for monetary compensation. One participant reported suspicions about the feedback provided, and six participants did not follow the instructions provided (e.g., did not cut all the shapes, cut the shapes in random order). We report here the analysis of data from the remaining 53 participants.

**Procedure.** This experiment employed a 3 (position of shape in the sequence: 1, 3, or 5) between-subjects design. Participants completed, in private rooms, an experiment assessing "how young adults perform on skills that they have (for the most part) underused since childhood." Participants' goal was to cut five identical shapes displayed on separate cards (see Figure 3). For this purpose, each participant received a large pair of scissors, an envelope and a stack of six cards displaying the same shape in different positions. The different positions ensured that participants did not attempt to cut several cards simultaneously. The written instructions asked participants to detach each card from the stack in the order presented (Image 1 first, etc.), carefully cut the shape displayed on the card at their own pace, place the cut image in the envelope provided, and then move on to the next card. To emphasize the goal sequence and task progress, information about the number of shapes completed so far was displayed at the top of each card (e.g., "Image 3 of 5"), and information about the number of shapes remaining was displayed at the bottom of the card (e.g., "2 more to go").

After each participant had completed the cutting task, the experimenter collected the envelope and asked the participant to wait. The experimenter returned a moment later with a result sheet for the participant. Participants did not know in advance that they would receive feedback on their performance, as this information might have influenced their performance for reasons beyond self-signaling concerns. The result sheet indicated "due to time constraints, the following feedback will cover only one of your shapes" and "no information will be provided for the other four shapes." Depending on the condition, participants received feedback on their first, third, or last shape. The result sheet informed them that their cutting performance for "Shape 1 (3 or 5) was compared with that of a prototypical average participant and was found to be above average." The feedback was positive, intentionally generic, and relative—rather than absolute—so that it would be more believable and interpreted positively, regardless of participants' awareness of their actual performance.

Next, participants read some general information linking cutting abilities to eye-hand coordination and fine motor skills (dexterity). Then, to assess the perceived diagnosticity of one's performance, we had participants rate their own (a) cutting skill level, (b) eye-hand coordination, and (c) dexterity, on 10-point scales (1 = *very low*; 10 = *very high*). The rationale of these measures is that feedback about more diagnostic actions should have a greater influence on self-appraisals. At the end of the experiment, the experimenter thanked and debriefed each participant.

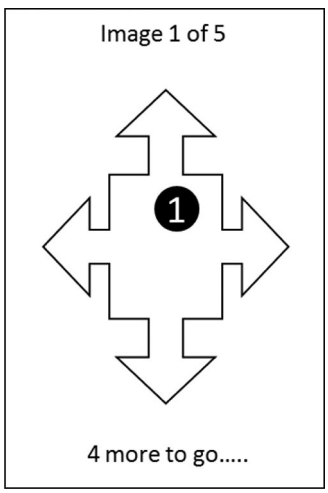


*Figure 3.* The shape used in the cutting shape task (Experiment 4).

## Results and Discussion

Two independent coders rated the cutting quality of all shapes by grouping the shapes into distinct categories based on how well the shapes were cut (e.g., how much participants cut corners, cut inside the line, or cut too far outside the line). The categories were then mapped onto a 10-point scale (1 = *poor*; 10 = *excellent*). A third independent coder resolved differences of 3 points or more between the first two coders' ratings by assigning the most accurate rating of the two to the shape (16 such discrepancies were resolved, out of 310 shapes). We averaged the ratings to form a cutting-quality index ($r$s ranged from .76 to .90).

A one-way within-subject ANOVA of the cutting-quality index yielded the predicted effect for position in the sequence, $F(4, 208) = 2.58$, $p < .05$. Further analysis revealed a quadratic contrast, $F(1, 52) = 6.87$, $p < .01$, indicating that participants cut shapes at the beginning and end better than shapes in the middle (Figure 4). For the rest of the analysis, we focused on Shapes 1, 3, and 5, for which we also had diagnosticity ratings. In congruence with the quadratic contrast, participants cut Shape 1 ($M = 7.70$, $SD = 1.55$) better than Shape 3 ($M = 7.13$, $SD = 1.92$), $t(52) = 2.65$, $p < .01$. They also cut Shape 5 ($M = 7.49$, $SD = 1.80$) better than Shape 3, $t(52) = 2.03$, $p < .05$. We found no difference between the cutting quality of the first and last shapes, $t(52) = 1.11$, $p = .27$.

In the domain of performance standards, we thus found that participants were more likely to literally cut corners in the middle of the sequence than at the beginning and end. We note that we replicated this pattern in a follow-up study that used 6 shapes instead of 5, $F(5, 95) = 5.75$, $p < .001$, with a similar quadratic contrast, $F(1, 19) = 20.62$, $p < .001$. We can thus conclude that the position of an action in a sequence (beginning vs. middle vs. end) rather than natural fluctuations in resource investment drives adherence to standards.

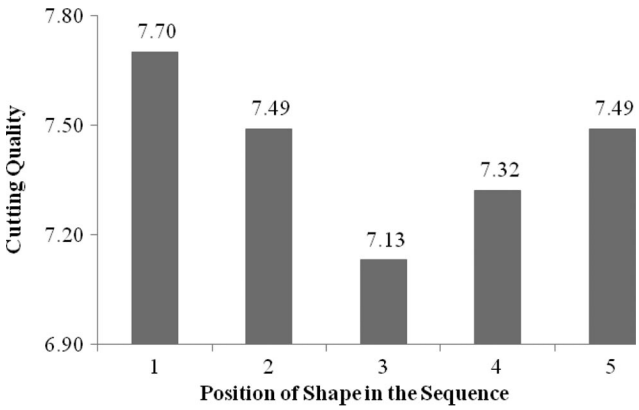


*Figure 4.* Average ratings of quality of cutting for each of the five shapes in the sequence (Experiment 4).

Next, we performed a between-subjects analysis of the effect of position on the perceived diagnosticity of actions. We combined self-appraisals of cutting skill, eye-hand coordination, and dexterity ($\alpha = .87$) to form a measure of diagnosticity. An ANOVA confirmed the predicted effect of position on diagnosticity, $F(2, 50) = 3.98$, $p < .05$. Participants who received positive feedback about the first and last shapes evaluated themselves as more skilled ($M = 8.14$, $SD = 0.88$ and $M = 8.20$, $SD = 1.36$, respectively) than did participants who received the same feedback about the third shape ($M = 6.83$, $SD = 2.31$), $t(22.1) = 2.24$, $p < .05$ (vs. first shape, equal variance not assumed); $t(34) = 2.17$, $p < .05$ (vs. last shape). Feedback about the first and last shape produced similar self-evaluations ($t < 1$, *ns*).

These results indicated that people see first and last actions in the course of goal pursuit as more diagnostic than middle actions when making self-inferences, in this case, about their own skill level. It is possible that people attributed their initial performance to "natural talent," whereas they attributed subsequent performance to practice or learning, assuming they believe their abilities in this particular context are malleable (Dweck, 2000; Dweck & Leggett, 1988). However, in that case, the perceived diagnosticity of the last (vs. middle) trial should not have been higher since it was more likely to reflect learning. Instead, we propose that the salience of beginning- and end-actions renders them more diagnostic. Interestingly, if participants were aware of their own tendency to relax performance standards in the middle, receiving positive performance feedback about their middle actions should have made them feel more skilled because they performed well even without trying very hard. We can thus assume that people are unaware of the tendency to relax standards in the middle.

## Mediation Analysis

To test whether perceived diagnosticity of actions mediates the effect of position (beginning vs. end vs. middle) on task performance, we first coded position such that higher numbers indicated beginning and end (Shapes 1 and 5, combined) versus middle (Shape 3). Results showed that position predicted both diagnosticity ($\beta = .37$), $t(51) = 2.85$, $p < .001$, and performance ($\beta = .31$), $t(51) = 2.35$, $p < .05$. Moreover, diagnosticity predicted performance ($\beta = .41$), $t(51) = 3.24$, $p < .01$. However, when both diagnosticity and position were included in the regression as predictors of performance, diagnosticity predicted performance ($\beta = .35$), $t(50) = 2.53$, $p < .05$, whereas position no longer predicted performance ($\beta = .18$), $t(50) = 1.36$, $p = .18$, suggesting diagnosticity mediated the effect of position on performance (Sobel $z = 2.14$, $p < .05$).

Importantly, because we measured the mediator (perceived diagnosticity) after the performance variable, we conducted another analysis that treated performance as the mediator and perceived diagnosticity as the dependent variable. The nonsignificant mediation results of this analysis confirmed that performance did not mediate the relationship between position and perceived diagnosticity, further supporting our hypothesis about the mediating role of perceived diagnosticity.

We find that the motivation to adhere to standards follows a u-shaped pattern because beginning and end (vs. middle) actions appear to have higher self-signaling value. We note that at the beginning of the task, the notion that the task measures cutting skills and performance might have been more active in participants' memory, which could have contributed to improved performance at beginning trials—by making performance standards more salient. However, such accessibility could not account for the increase in adherence to standards toward the end of the task, when the temporal distance from the presentation of these standards was largest. In addition, the greater accessibility of standards at the beginning could not explain the effect of position on perceived diagnosticity. Instead, we argue that greater adherence to standards is the result of the greater perceived diagnosticity of actions.

We propose that people's motivation to adhere to their standards operates independently of their motivation to reach a goal's end state; that is, doing it right is a different aspect of motivation than getting it done. To demonstrate this point, we conducted our final study that examined whether, within the same self-regulatory task, the outcome-focused motivation to complete a goal would increase monotonically (as in goal gradient studies, e.g., Hull, 1932), whereas the motivation to adhere to standards would follow a u-shaped pattern.

## Experiment 5: Doing It Right Versus Getting It Done

To distinguish between the patterns of means-focused motivation to adhere to standards and outcome-focused motivation to reach a goal's end state, we designed a task in which we could measure adherence to standards separately from motivation to complete the task. Specifically, participants completed a lexical task consisting of a series of trials, and we assessed the extent to which they (a) applied themselves to perform well on each trial and (b) were eager to finish the task by moving more quickly from one trial to the next. We predicted that participants would perform better on the first and last trials than on middle trials because first and last trials appear highly diagnostic. We further predicted that they would be more motivated to finish the task as they approached the last trial—that is, they would take shorter breaks between trials because the psychophysics of goal pursuit render each subsequent trial seemingly more impactful on goal attainment (Brown, 1948; Förster et al., 1998; Hull, 1932; Kivetz et al., 2006).

## Method

**Participants.** Fifty eight students from various universities in the greater Chicago area (23 men, 35 women) participated in the

experiment for monetary compensation. All students were native English speakers. We excluded two participants from our analysis, based on their total scores in the task. Out of the 42 word pairs presented in the task, these two participants solved 0 and 2, which was more than 3 standard deviations under the mean scores ($M$ = 34.9, $SD$ = 9.84; see Bargh & Chartrand, 2000).

**Procedure.** This experiment employed a three (position of trial in the sequence: 1, 4, or 7) within-subjects design. Participants completed in private rooms an experiment assessing "attention to detail, observation, and verbal skills." Participants' goal was to complete a series of seven trials in which they were presented with six pairs of words. For each pair, they had to spot and identify the word with the fewest letters (i.e., the shortest word). For example, for the pair *asparagus–positivity*, the word with the fewest letters was asparagus. Participants read that they could "count or estimate the number of letters in each word" and were made aware that "sometimes words might appear shorter than they actually are because some letters take less space than others." Therefore, participants who wanted to apply themselves might be more likely to use a counting (vs. estimation) strategy in order to get the correct answer.

Words were between nine and 12 letters long, and each pair consisted of words that were only one letter apart (e.g., a nine-letter vs. a 10-letter word). Before starting the main study, participants completed a short practice task (featuring two pairs of words). Each of the seven trials in the main study consisted of six pairs of words. We counterbalanced the order of Trials 1(first), 4 (middle) and 7 (last) to control for any trial-specific effects and focused our analysis of performance on these three trials. As in previous studies, we ensured that participants noted their progress in the task by displaying a progress chart before each trial and numbering each trial (e.g., Trial 4 of 6).

We measured two distinct aspects of motivation. First, we measured the motivation to adhere to one's standards by assessing trial performance. For this purpose, we computed the number of pairs (up to 6) for which participants correctly identified the shorter word. Second, we measured participants' eagerness to reach the goal's end state by recording the speed with which they moved from one trial to the next. The study instructions asked participants to "move at [their] own pace and feel free to take short breaks between trials if needed," and each of the six intermediate screens between trials reiterated this information and featured progress information. The amount of time participants spent on each of these intermediate screens (intertrial time) constituted our measure of eagerness to finish the series of trials (see also Custers & Aarts, 2005). We chose to measure intertrial times because unlike the amount of time a participant spends working on each trial (trial time), shorter intertrial times reflected eagerness to complete the task, and intertrial times were less likely to be influenced by learning or practice. After participants completed the seven trials, we assessed their affect and attitude toward each of the three critical trials (1, 4, and 7) by asking them about the extent to which they found each of these trials pleasant, interesting, and difficult (1 = *not at all*; 7 = *very much*). They were then debriefed and dismissed.

## Results and Discussion

We focused our analysis of performance on Trials 1 (beginning), 4 (middle) and 7 (end) because the order of these three trials was counterbalanced to control for any trial- or order-specific effects, whereas the remaining trials (2, 3, 5, and 6) appeared in the same position in all conditions. We found no effect for counterbalancing. We conducted a one-way ANOVA of performance on the three target trials, controlling for overall level of performance—which was the sum of scores for the four nontarget trials (2, 3, 5, and 6). We included this overall performance measure as a covariate because it served as a proxy for participants' ability and distinguished between participants who were more versus less skilled at the task. This analysis yielded an effect for position in the sequence, $F(2, 108) = 3.24$, $p < .05$, and revealed the predicted quadratic contrast, $F(1, 54) = 7.74$, $p < .01$. Participants found more correct answers for the first and last trials ($M = 5.21$, $SD = 1.16$ and $M = 5.20$, $SD = 1.21$, respectively) than for the middle trial ($M = 5.05$, $SD = 1.52$), $F(1, 54) = 5.04$, $p < .05$ (vs. first trial); $F(1, 54) = 5.45$, $p < .05$ (vs. last trial). However, there was no difference in performance between the first and last trials ($F < 1$, *ns*). These results show that participants adhered more closely to performance standards at the beginning and end (vs. middle) of goal pursuit.

Next, we performed an ANOVA of intertrial times: the amount of time participants spent on each of the six intermediate screens between the seven trials. We excluded intertrial times that were over 3 standard deviations above the means of intertrials times (Bargh & Chartrand, 2000). Mauchly's test indicated the assumptions of sphericity had been violated, $\chi^2(14) = 88.46$, $p < .001$; therefore, we corrected degrees of freedom using Greenhouse-Geisser estimates of sphericity ($\varepsilon = .57$). The results revealed an effect for position, $F(2.87, 132.12) = 18.63$, $p < .001$, and the predicted linear contrast, indicating that intertrial times decreased as participants moved from the beginning to the end of the seven-trial sequence, $F(1, 46) = 35.46$, $p < .001$ (see Figure 5). Thus, participants became more eager to finish the task as they approached the last trial.

We also note that there were no correlations between performance on each of the three target trials (1, 4, and 7) and the corresponding intertrial time(s) measured before and/or after the task ($-.18 < r\text{s} < .13$, *ns*). The lack of correlation between these two variables further supports our contention that they measure two independent aspects of motivation.

Finally, ANOVA on participants' affect and attitudes toward each of the three critical trials (1, 4, and 7), measured at the end of the task, revealed no effect of position on perceptions that the trials

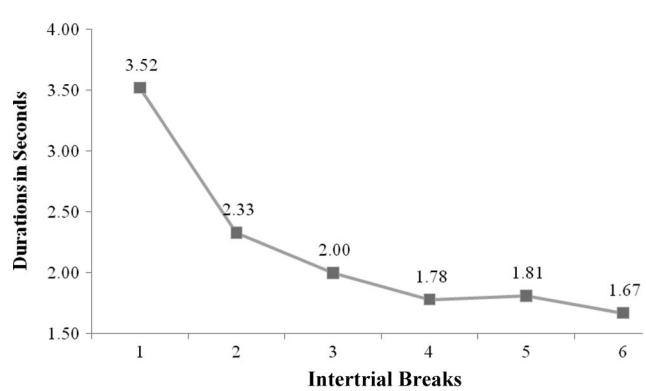


*Figure 5.* Intertrial break durations (in seconds) as a function of progress in the task (Experiment 5).

were pleasant, $F(1.72, 94.62) = .81$, *ns*, or difficult, $F(1.74, 95.5) = 1.17$, *ns*—note that degrees of freedom were corrected using Huynh-Feldt estimates of sphericity ($\varepsilon s > .75$), since Mauchly's test indicated the assumptions of sphericity had been violated. However, position influenced perceptions that the trials were interesting, $F(2, 110) = 8.63$, $p < .001$. The first trial ($M = 4.3$, $SD = 1.99$) appeared more interesting than both the middle trial ($M = 3.82$, $SD = 1.89$) and the last trial ($M = 3.82$, $SD = 1.96$), with $t(55) = 3.58$, $p = .001$, and $t(55) = 3.86$, $p < .001$, respectively. Middle and last trials were thought to be equally interesting ($t < 1$, *ns*). This pattern could not account for the u-shaped effect of position on performance discussed above. Specifically, low performance in the middle could not be the result of boredom or loss of interest in the middle of the task, compared with the beginning and end.

This experiment uncovers two distinct aspects of motivation in the course of goal pursuit, showing that these motivations follow different patterns and are captured by different types of measures. First, the motivation to adhere to performance standards follows a u-shaped pattern, such that participants apply themselves more on tasks at the beginning and end (vs. middle). Second, the motivation to complete the task increases monotonically, as previously indicated by goal-gradient research.

## General Discussion

In the course of pursuing goals that require the completion of a sequence of actions, people can decide how closely to adhere to their personal standards for each action, based on whether the benefits of relaxing these standards outweigh the costs. One such cost is the negative impact on self-image because relaxing standards can signal to a person that he or she has low standards (Batson et al., 1997, 1999; Mazar et al., 2008). We find that these self-signaling concerns fluctuate with the perceived diagnosticity of actions over the course of goal pursuit, leading people to adhere more closely to their standards at the beginning and end of goal pursuit than in the middle.

We observe this pattern of behavior for public and private transgressions in the context of ethical, religious, and performance standards. Across our experiments, participants were more likely to behave unethically in the middle of goal pursuit than at the beginning and end by falsely reporting the favorable outcome of a coin flip (Experiment 1) and by taking advantage of an experimenter's "forgetfulness" to get undeserved credit (Experiment 2). Similar patterns emerged for adherence to religious and performance standards: in the middle, participants skipped religious rituals (Experiment 3), literally cut corners (Experiment 4), and opted for estimating rather than counting (Experiment 5).

We further find that judgments of diagnosticity underlie our effects, such that people perceive actions at the beginning and end (vs. middle) of a sequence of actions as a better signal of their true character. We explore this underlying process using several methodological approaches. First, we show that the degree of self-relevance of the standard under consideration moderates the effect of position in the sequence. In Experiment 3A, participants' degree of religiousness moderated their pattern of standard adherence over the course of pursuing religious rituals. Second, we show that people's perception of the differential diagnosticity of their performance at the beginning and end (vs. middle) of a task mediates the effect of position in the sequence on the actual quality of their work (Experiment 4). Thus, when people expect their performance to be highly representative of their abilities, they work harder and perform better. Third, we demonstrate that this pattern of differential diagnosticity extends to judgment of others, such that people judge others more harshly for religious transgressions committed at the beginning and end of goal pursuit than for those committed in the middle (Experiment 3B).

### Implications for Motivation and Self-Awareness Theories

These results have implications for existing motivation theory. We distinguish between two dimensions of motivation in the course of pursuing a goal: (a) the motivation to achieve the focal goal or reduce the discrepancy between current and end states and (b) the motivation to do it right in order to maintain a positive self-image in the process of reaching the end state (Touré-Tillery & Fishbach, 2011). The motivation to approach the focal goal (outcome-focused motivation) determines the amount of effort and persistence channeled toward goal pursuit, including how fast people work and how eager they are to reach the goal. Previous research documented several different patterns for this aspect of motivation, including a goal-gradient pattern of increasing effort and persistence as distance from the goal decreases (Brown, 1948; Förster et al., 1998; Hull, 1932; Kivetz et al., 2006), and a u-shaped pattern of higher resource investment at the initial and end (vs. middle) stages of goal pursuit corresponding to the perceived marginal value of progress at each stage (Bonezzi et al., 2011). By contrast, the motivation to use proper means of goal pursuit and maintain a positive self-image primarily influences how closely people follow their standards, that is, how concerned they are with performing actions the "right" way (means-focused motivation; see also Higgins, Idson, Freitas, Spiegel, & Molden, 2003). We document a u-shaped pattern of adherence to standards, corresponding to the pattern of mean-focused motivation.

These two dimensions of motivations can be in conflict if they both happen to be high at a particular stage of goal pursuit. When, for example, outcome-focused motivation follows a goal-gradient pattern, we would expect a greater conflict toward the end of goal pursuit than at the beginning or in the middle. Indeed, at the beginning of goal pursuit, the u-shaped motivation to maintain a positive self-image by doing things the right way should dominate goal attainment motives, whereas in the middle, the opposite should be true. Then, as people approach the "finish line," both motivations should increase in strength, creating a stronger conflict between reaching the end state and doing things properly. Although the present research does not assess the magnitude of this conflict, the presence of these conflicting forces at the end of goal pursuit could potentially explain why adherence to standards sometimes seemed directionally greater at the beginning than at the end of goal pursuit (see Experiments 1 and 3).

The present results also have implications for research on self-awareness. In some of our experiments, participants might have been more self-aware at the beginning and end of the task sequence than in the middle, possibly because of the salience and greater perceived diagnosticity of these actions. This increased self-awareness might in turn have contributed to the standard adherence patterns of behavior observed. Indeed, according to Duval and Wicklund's (1972) self-awareness theory, self-focused

attention makes people more conscious of their attitudes and beliefs (Goukens, Dewitte, & Warlop, 2009) and leads to greater adherence to personal standards of behavior (Carver, 1975). Although, the present research does not delve into the specific connections between salience, diagnosticity, and self-awareness, we believe such links might exist under certain circumstances. For instance, in some cases, self-awareness may very well be an additional mediating factor of the effect of position and/or perceived diagnosticity on adherence to standards. Such a role of self-awareness would be consistent with our theory in some contexts. However, we show that diagnosticity considerations also influence interpersonal judgments, such that people are more likely to make inferences about another person's character from her beginning and end actions than from her middle actions. This impact of position in a sequence on inferences of others' traits cannot be mediated by self-awareness, which leads us to believe in the existence of a link between position, perceived diagnosticity and adherence to standards that does not require changes in self-awareness. Future research could explore these links between diagnosticity and self-awareness more thoroughly.

## The Subjectivity of Position in a Sequence

Our experiments had sequences of varying lengths and showed that for longer sequences, the beginning and end often included more actions than merely the first and last, whereas for shorter sequences, the beginning and end were roughly represented by the first and last actions. This finding suggests the position of an action in a sequence is often subjective and depends on a person's perception of relative location in the sequence rather than an objective or absolute location. For example, a third action could be seen as part of the beginning of a 50-action sequence, the middle of a five-action sequence, or the end of a three-action sequence. Performance on this third action in terms of how much people would adhere to their standards would then depend on the overall length of the sequence and on people's perceptions of their locations in the sequence. If people believe (or are led to believe) they are at the beginning or approaching the end of a sequence of actions, they should act more in line with their standards than if they believe they are in the middle.

Combined with our findings, the understanding that perceptions of beginning, middle, and end are malleable has important practical implications for designing goal sequences or framing actions in ways that promote standard adherence. For example, by presenting an action (e.g., project, task, or exam) as the first or the last of a constructed sequence of actions, decision makers in management or education could increase employees' and students' adherence to standards—without directly mentioning words such as "integrity" or "honor code." Moreover, future research could investigate such framing effects and determine whether merely priming the idea of beginning versus middle versus end would elicit judgmental and behavioral responses similar to the ones documented in the present research. Finally, by framing the beginnings and ends of sequences as encompassing more tasks or dividing long sequences of tasks into shorter sequences with several beginnings and ends, task designers could also increase the number of points of standard adherence along a sequence. Our research suggests that in managerial and educational contexts, such simple positioning of actions could substantially raise standard adherence by increasing not only the quality of outputs but also the likelihood that these outputs are produced ethically.

## References


Allingham, M. G., & Sandmo, A. (1972). Income tax evasion: A theoretical analysis. *Journal of Public Economics, 1,* 323–338. doi:10.1016/0047-2727(72)90010-2

Bandura, A. (1977). *Social learning theory.* Englewood Cliffs, NJ: Prentice Hall.

Bargh, J. A., & Chartrand, T. L. (2000). The mind in the middle: A practical guide to priming and automaticity research. In H. T. Reis & C. M. Judd (Eds.), *Handbook of research methods in social and personality psychology* (pp. 253–285). New York, NY: Cambridge University Press.

Batson, C. D., Kobrynowicz, D., Dinnerstein, J. L., Kampf, H. C., & Wilson, A. D. (1997). In a very different voice: Unmasking moral hypocrisy. *Journal of Personality and Social Psychology, 72,* 1335–1348. doi:10.1037/0022-3514.72.6.1335

Batson, C. D., Thompson, E. R., Seuferling, G., Whitney, H., & Strongman, J. A. (1999). Moral hypocrisy: Appearing moral to oneself without being so. *Journal of Personality and Social Psychology, 77,* 525–537. doi:10.1037/0022-3514.77.3.525

Baumeister, R. F., & Newman, L. S. (1994). Self-regulation of cognitive inference and decision-processes. *Personality and Social Psychology Bulletin, 20,* 3–19. doi:10.1177/0146167294201001

Becker, G. S. (1968). Crime and punishment—Economic approach. *Journal of Political Economy, 76,* 169–217. doi:10.1086/259394

Bem, D. J. (1972). Self-perception theory. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 6, pp. 1–62). New York, NY: Academic Press.

Bodner, R., & Prelec, D. (1996). The emergence of private rules in a self-signalling model. *International Journal of Psychology, 31,* 3652–3653.

Bonezzi, A., Brendl, M., & De Angelis, M. (2011). Stuck in the middle: The psychophysics of goal pursuit. *Psychological Science, 22,* 607–612. doi:10.1177/0956797611404899

Brown, J. S. (1948). Gradients of approach and avoidance responses and their relation to level of motivation. *Journal of Comparative and Physiological Psychology, 41,* 450–465. doi:10.1037/h0055463

Burton, R. V., & Kunce, L. (1995). behavioral models of moral development: A brief history and integration. In W. M. Kurtines & J. L. Gewirtz (Eds.), *Moral development: An introduction* (pp. 141–171). Boston, MA: Allyn & Bacon.

Carver, C. S. (1975). Physical aggression as a function of objective self-awareness and attitudes toward punishment. *Journal of Experimental Social Psychology, 11,* 510–519. doi:10.1016/0022-1031(75)90002-5

Custers, R., & Aarts, H. (2005). Positive affect as implicit motivator: On the nonconscious operation of behavioral goals. *Journal of Personality and Social Psychology, 89,* 129–142. doi:10.1037/0022-3514.89.2.129

Duval, S., & Wicklund, R. A. (1972). *A theory of objective self-awareness.* New York, NY: Academic Press.

Dweck, C. S. (2000). *Self-theories: Their role in motivation, personality and development.* Philadelphia, PA: Psychology Press.

Dweck, C. S., & Leggett, E. L. (1988). A social cognitive approach to motivation and personality. *Psychological Review, 95,* 256–273. doi:10.1037/0033-295X.95.2.256

Eidelman, S., & Biernat, M. (2007). Getting more from success: Standard raising as esteem maintenance. *Journal of Personality and Social Psychology, 92,* 759–774. doi:10.1037/0022-3514.92.5.759

Eisenberg, N., & Shell, R. (1986). Pro-social moral judgment and behavior in children—The mediating role of cost. *Personality and Social Psychology Bulletin, 12,* 426–433. doi:10.1177/0146167286124005

Festinger, L. A. (1957). *A theory of cognitive dissonance.* Stanford, CA: Stanford University Press.

Fishbach, A., Dhar, R., & Zhang, Y. (2006). Subgoals as substitutes or complements: The role of goal accessibility. *Journal of Personality and Social Psychology, 91,* 232–242. doi:10.1037/0022-3514.91.2.232

Fishbach, A., Henderson, D. H., & Koo, M. (2011). Pursuing goals with others: Group identification and motivation resulting from things done versus things left undone. *Journal of Experimental Psychology: General, 140,* 520–534.

Fishbach, A., Zhang, Y., & Koo, M. (2009). The dynamics of self-regulation. *European Review of Social Psychology, 20,* 315–344. doi:10.1080/10463280903275375

Förster, J., Higgins, E. T., & Idson, L. C. (1998). Approach and avoidance strength during goal attainment: regulatory focus and the "goal looms larger" effect. *Journal of Personality and Social Psychology, 75,* 1115–1131. doi:10.1037/0022-3514.75.5.1115

Gino, F., Ayal, S., & Ariely, D. (2009). Contagion and differentiation in unethical behavior: the effect of one bad apple on the barrel. *Psychological Science, 20,* 393–398. doi:10.1111/j.1467-9280.2009.02306.x

Gino, F., & Pierce, L. (2009). Dishonesty in the name of equity. *Psychological Science, 20,* 1153–1160. doi:10.1111/j.1467-9280.2009.02421.x

Gollwitzer, P. M., Wicklund, R. A., & Hilton, J. L. (1982). Admission of failure and symbolic self-completion: Extending Lewinian theory. *Journal of Personality and Social Psychology, 43,* 358–371. doi:10.1037/0022-3514.43.2.358

Goukens, C., Dewitte, S., & Warlop, L. (2009). Me, myself, and my choices: The influence of private self-awareness on choice. *Journal of Marketing Research, 46,* 682–692. doi:10.1509/jmkr.46.5.682

Greene, R. L. (1986). Sources of recency effects in free-recall. *Psychological Bulletin, 99,* 221–228. doi:10.1037/0033-2909.99.2.221

Greenwald, A. G., & Breckler, S. J. (1985). To whom is the self presented? In B. R. Schlenker (Ed.), *The self and social life* (pp. 126–145). New York, NY: McGraw-Hill.

Heath, C., Larrick, R. P., & Wu, G. (1999). Goals as reference points. *Cognitive Psychology, 38,* 79–109. doi:10.1006/cogp.1998.0708

Higgins, E. T., & Brendl, C. M. (1995). Accessibility and applicability: Some activation rules influencing judgment. *Journal of Experimental Social Psychology, 31,* 218–243. doi:10.1006/jesp.1995.1011

Higgins, E. T., Idson, L. C., Freitas, A. L., Spiegel, S., & Molden, D. C. (2003). Transfer of value from fit. *Journal of Personality and Social Psychology, 84,* 1140–1153. doi:10.1037/0022-3514.84.6.1140

Hoffman, M. L. (1977). Moral internalization: Current theory and research. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 10, pp. 86–135). New York, NY: Academic Press.

Hogan, R., Jones, W. H., & Cheek, J. M. (1985). Socioanalytic theory: An alternative to armadillo psychology. In B. R. Schlenker (Ed.), *The self and social life* (pp. 175–198). New York, NY: McGraw-Hill.

Hull, C. L. (1932). The goal gradient hypothesis and maze learning. *Psychological Review, 39,* 25–43. doi:10.1037/h0072640

Kivetz, R., Urminsky, O., & Zheng, Y. H. (2006). The goal-gradient hypothesis resurrected: Purchase acceleration, illusionary goal progress, and customer retention. *Journal of Marketing Research, 43,* 39–58. doi:10.1509/jmkr.43.1.39

Koo, M., & Fishbach, A. (2008). Dynamics of self-regulation: How (un-)accomplished goal actions affect motivation. *Journal of Personality and Social Psychology, 94,* 183–195. doi:10.1037/0022-3514.94.2.183

Leary, M. R., & Kowalski, R. M. (1990). Impression management: A literature-review and two-component model. *Psychological Bulletin, 107,* 34–47. doi:10.1037/0033-2909.107.1.34

Lewicki, R. J. (1984). Lying and deception: A behavioral model. In M. H. Bazerman & R. J. Lewicki (Eds.), *Negotiation in organizations* (pp. 68–90). Beverly Hills, CA: Sage.

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. doi:10.1509/jmkr.45.6.633

Mead, N. L., Baumeister, R. F., Gino, F., Schweitzer, M. E., & Ariely, D. (2009). Too tired to tell the truth: Self-control resource depletion and dishonesty. *Journal of Experimental Social Psychology, 45,* 594–597. doi:10.1016/j.jesp.2009.02.004

Monin, B., & Miller, D. T. (2001). Moral credentials and the expression of prejudice. *Journal of Personality and Social Psychology, 81,* 33–43. doi:10.1037/0022-3514.81.1.33

Murdock, B. B. (1960). The distinctiveness of stimuli. *Psychological Review, 67,* 16–31. doi:10.1037/h0042382

Prelec, D., & Bodner, R. (2003). Self-signaling and self-control. In G. Loewenstein, D. Read, & R. F. Baumeister (Eds.), *Time and decision: Economic and psychological perspectives on intertemporal choice* (pp. 277–298). New York, NY: Russell Sage Foundation.

Pryor, J. B., & Kriss, M. (1977). Cognitive dynamics of salience in attribution process. *Journal of Personality and Social Psychology, 35,* 49–55. doi:10.1037/0022-3514.35.1.49

Sachdeva, S., Iliev, R., & Medin, D. L. (2009). Sinning saints and saintly sinners: The paradox of moral self-regulation. *Psychological Science, 20,* 523–528. doi:10.1111/j.1467-9280.2009.02326.x

Schlenker, B. R. (1985). Identity and self-identification. In B. R. Schlenker (Ed.), *The self and social life* (pp. 65–99). New York, NY: Mc-Graw-Hill.

Skowronski, J. J., & Carlston, D. E. (1987). Social judgment and social memory: The role of cue diagnosticity in negativity, positivity, and extremity biases. *Journal of Personality and Social Psychology, 52,* 689–699. doi:10.1037/0022-3514.52.4.689

Skowronski, J. J., & Carlston, D. E. (1989). Negativity and extremity biases in impression-formation: A review of explanations. *Psychological Bulletin, 105,* 131–142. doi:10.1037/0033-2909.105.1.131

Steele, C. M. (1988). The psychology of self-affirmation: Sustaining the integrity of the self. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 21, pp. 261–302). New York, NY: Academic Press.

Swann, W. B., & Read, S. J. (1981). Acquiring self-knowledge: The search for feedback that fits. *Journal of Personality and Social Psychology, 41,* 1119–1128. doi:10.1037/0022-3514.41.6.1119

Taylor, S. E., & Fiske, S. T. (1978). Salience, attention, and attribution: Top of the head phenomena. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 11, pp. 249–288). New York, NY: Academic Press.

Taylor, S. E., Fiske, S. T., Etcoff, N. L., & Ruderman, A. J. (1978). Categorical and contextual bases of person memory and stereotyping. *Journal of Personality and Social Psychology, 36,* 778–793. doi:10.1037/0022-3514.36.7.778

Touré-Tillery, M., & Fishbach, A. (2011). The course of motivation. *Journal of Consumer Psychology*, *21*, 414–483.

Vohs, K. D., & Schooler, J. W. (2008). The value of believing in free will: Encouraging a belief in determinism increases cheating. *Psychological Science, 19,* 49–54. doi:10.1111/j.1467-9280.2008.02045.x

Wooldridge, J. M. (2002). *Econometric analysis of cross section and panel data.* Cambridge, MA: MIT Press.

Wright, D. (1971). *The psychology of moral behavior.* Baltimore, MD: Penguin Books.

Zhong, C. B., Bohns, V. K., & Gino, F. (2010). Good lamps are the best police: Darkness increases dishonesty and self-interested behavior. *Psychological Science, 21,* 311–314. doi:10.1177/0956797609360754

Received March 6, 2011
Revision received August 1, 2011
Accepted September 12, 2011 ■