Self-Affirmation Can Enable Goal Disengagement

Kathleen D. Vohs
University of Minnesota

Brandon J. Schmeichel
Texas A&M University

Ji Kyung Park
University of Delaware

Much research has shown that after being self-affirmed, people respond to challenges in healthy, productive ways, including better task performance. The current research demonstrates that self-affirmation can also deflate motivation and performance, a pattern consistent with goal disengagement. We posited that being self-affirmed and then attempting but failing at a task would lead people to retreat from the goal. In support of this hypothesis, 4 experiments found that the combination of self-affirmation and the experience of failure led to demotivation and effort reduction. Experiment 1 found that self-affirmed participants, more so than nonaffirmed participants, reported being open to goal disengagement. Experiment 2 found that affirming core values before trying a task beset with failure reduced task motivation and performance. Experiment 3 demonstrated the robustness of the effect and found that failure on one task reduced motivation and performance on a new but related task. Experiment 4 revealed that being self-affirmed and experiencing failure caused participants to feel less capable of pursuing their goals, which produced poorer performance. These findings suggest that affirming the self can lead people to internalize the implications of failure, which in turn leads to goal disengagement.

Keywords: self-affirmation, goal disengagement, failure, motivation, performance

To reflect on what matters in life is a sophisticated mental act that can set one’s sights on new horizons or remind oneself of the preciousness of the human condition. One process that makes use of life reflections is self-affirmation, which confers a wellspring of benefits to those who practice it (Harris, Mayle, Mabott, & Napper, 2007; Martens, Johns, Greenberg, & Schimel, 2006; Miyake et al., 2010; Spencer, Fein, & Lomore, 2001). In contrast to most previous research, the current experiments tested the hypothesis that engaging in self-affirmation can dampen performance and motivation. We propose that self-affirmation combined with the experience of failure causes goal disengagement, or withdrawal of effort and motivation from obstinate focal goals.

Self-affirmation refers to cognitive events that support a global sense of self-integrity (Steele, 1988). Self-affirmation manipulations typically involve reflecting upon or expressing conviction in valued aspects of self or life. Research indicates that affirming core values helps to bolster the legitimacy of the self and promotes reacting with equanimity to threats to self-regard. For instance, when told of potential problems that might befall them, people who have affirmed the self show reduced defensiveness and vow to make positive behavioral changes more than nonaffirmed persons do (Harris et al., 2007; Harris & Napper, 2005; Reed & Aspinwall, 1998; Sherman & Cohen, 2002; Sherman, Nelson, & Steele, 2000).

We reasoned that self-affirmation may not always improve performance and motivation. People who have engaged in self-affirmation tend to view life from a different and, arguably, more dispassionate perspective. What self-affirmation does is to get people to accept information about their personal flaws as credible and view those flaws as plausible causes of future problems. In past work, self-affirmed people have been shown to react to these humbling realizations by intending to bring the self in line with more virtuous standards (Sherman & Cohen, 2006). In the current work, we proposed that affirming the self may contribute to another goal-regulation process, one rooted in reduced motivation. That process is goal disengagement, the process by which people abandon attempts to achieve a current goal that is blocked or recalcitrant. According to our model, self-affirmation allows information that one is not succeeding or likely to succeed to be heard and heeded. This means that people will buy into such information by altering probabilities of success and perceptions of the self’s abilities, which ultimately will lead to reduced motivation. In short, we predicted that compared to nonaffirmed persons, self-affirmed persons will react to failure (i.e., the experience of striving but not succeeding at a focal task) by disengaging from the goal.

Self-Threat and Self-Affirmation

Many studies have found that self-affirmation reduces defensive responses to threatening information. For example, after receiving...
negative feedback on an intelligence test, nonaffirmed individuals sought downward (and presumably palliative) social comparisons. By contrast, among those who had affirmed a core value, negative feedback led to upward social comparisons that provided opportunities to learn from star performers (Spencer et al., 2001). Another set of studies found that affirming core values reduced the self-serving bias, which is the ego-protective tendency to blame external factors for personal failures but to credit the self for successes (Sherman & Kim, 2005). Self-affirmation can also offset the influence of self-relevant stereotypes that harm performance on intelligence tests (e.g., pertaining to women or African Americans) and boost test scores (Cohen, Garcia, Purdie-Vaughns, Apfel, & Brzustoski, 2009; Martens et al., 2006; Miyake et al., 2010).

Self-affirmed people also take initial steps to stop engaging in risky behaviors. For example, they are more likely than nonaffirmed persons to assert intentions to change (Harris et al., 2007; Harris & Napper, 2005; Reed & Aspinwall, 1998), purchase condoms (Sherman et al., 2000), and obtain health-risk informational brochures (Armitage, Harris, Hepton, & Napper, 2008). Hence, previous research has indicated that self-affirmation allows people to respond to failure, negative feedback, threatening stereotypes, and health concerns by forging new behaviors that promise to reduce the potential for future damage.

Self-affirmation seems to allow threatening information about the self to be heard and believed. Various self-affirmation exercises have been shown to encourage acceptance of ominous health messages in that affirmed people saw themselves as having relatively higher odds for contracting illness or disease (Epton & Harris, 2008; Harris & Napper, 2005; Sherman & Cohen, 2006). Similarly, Klein and Harris (2009) found that female alcohol drinkers who had vs. had not been self-affirmed devoted more attention to information linking alcohol consumption to breast cancer. Acknowledging that one is personally vulnerable to potentially negative outcomes helps to explain why self-affirmation leads to virtuous intentions to change risky or unhealthy behaviors (Klein, Harris, Ferrer, & Zajac, 2011).

Hence, an open and realistic acknowledgment of one’s personal characteristics (perhaps especially one’s flaws) is central to how self-affirmation helps to regulate goal-directed behavior. We view such sensitivity as key to understanding why people who have been self-affirmed may react to failure with goal disengagement.

**Goal Failure and Disengagement**

An unfortunate reality is that to attain many valuable goals in life, one must battle failure. Another unfortunate reality is that many such battles are met with only more failure. People may start the process of disengagement when they run into impediments in goal pursuit that prompt them to evaluate the likelihood of ever achieving the goal (Carver & Scheier, 1990; Nesse, 2000). Although giving up on a goal can be viewed as an undesirable outcome, it can also be beneficial. Research has found that adolescent girls who disengaged from unattainable goals showed drops in an immunomarker of the body's inflammatory responses (Miller & Wrosch, 2007). Likewise, adolescent girls who shifted efforts away from failing goals and toward promising ones experienced a decline in depressive symptoms (Wrosch & Miller, 2009). Therefore, some evidence suggests that goal disengagement can lead to improvements in physiological and psychological well-being.

But people may be reluctant to disengage from goals, particularly when the goals are salient or valuable, or when disengagement carries negative implications for the self (Brockner et al., 1986). Consistent with this view, research has found that negative feedback and other threats to self-regard heighten the tendency to commit additional time and resources to losing courses of action (e.g., Zhang & Baumeister, 2006). If the defensive response to threatened self-regard is stubborn persistence at intransigent goals, then self-affirmation—which is known to reduce defensive responding to threat—may facilitate goal disengagement.

**The Importance of Self-Efficacy Perceptions**

Due to self-affirmation’s ability to lower defensiveness and increase openness to negative information about the self, we reasoned that being self-affirmed before trying and failing at a task could lead to lower perceptions of self-efficacy. If that is correct, then downward revisions of self-efficacy may account for the proposed effects of self-affirmation on goal disengagement.

Self-perceptions are an important component of self-affirmation’s effects on motivation. Most previous research has observed that self-affirmation exercises boost perceptions of self-efficacy. A study on the dangers of caffeine consumption, for instance, found that participants who thought about past examples of their own considerate and kind behaviors felt more assured that they could control their caffeine consumption, compared to counterparts who had not thought about their past considerate behaviors (Reed & Aspinwall, 1998). This same manipulation was used in a campaign to improve fruit and vegetable intake among women. Self-affirmed participants agreed more strongly with statements such as “I know for sure that I could adhere to eating at least 5 portions of fruit and vegetables each day” (Epton & Harris, 2008). Another study instructed cigarette smokers to think about their values, whereas others were not given this instruction. Then all participants viewed antismoking public service announcements. Affirmed smokers, as compared to nonaffirmed smokers, showed an uptick in self-efficacy as measured by the statement “How sure are you that you could stay away from cigarettes if you really wanted to?” (Zhao & Nan, 2010). Thus, self-affirmation can elicit confidence in one’s ability to meet goals.

We too hypothesized that self-affirmation would change perceptions of self-efficacy. Like the studies reviewed, we predicted that self-affirmation would boost self-efficacy in the absence of failure. In the context of a salient failure experience, however, we hypothesized that self-affirmation would invite a downturn in self-efficacy.

These predictions are grounded in one of the clearest conclusions from the self-affirmation literature: People who have recently engaged in self-affirmation, more so than those who have not, acknowledge their personal vulnerabilities and appreciate that bad outcomes may lie ahead (Klein et al., 2011; Sherman & Cohen, 2006; Tesser, 2000). In the context of pursuing a goal beset with setbacks, this dispassionate perspective may help people to realize that they are not capable of achieving the focal goal. We therefore posited that because self-affirmed persons are open to negative information about the self, failure may induce the belief
that one is not well equipped to reach the goal, a perception that could precipitate goal disengagement.

This hypothesis converges with research from the goal-dissatisfaction literature insofar as self-perceptions have been implicated in the act of distancing oneself from difficult goals. Kukla (1972) and others (e.g., Wright & Dismukes, 1995) have observed that self-efficacy and motivation relate to one another in the sense that if people perceive that they do not possess the capacity to obtain their goals, they stop trying as hard as they would otherwise. Empirical support for these notions was provided by Vancouver, More, and Yoder (2008) in an experiment that manipulated self-efficacy and then measured the effort participants committed to a task. The lower participants’ perceptions of self-efficacy, the fewer resources (meaning the less motivation) they put toward achieving a goal.

In the current research, we predicted that people who self-affirmed and then failed to meet a goal would come to see themselves as less able to achieve goals generally. This, in turn, would result in diminished effort and motivation to reach the current goal. In essence, then, we posited that the perception that one may not reach one’s goal serves as a signal to self-affirmed persons to begin the process of goal disengagement.

The Present Experiments

We proposed that goal disengagement can result from feeling self-affirmed in combination with goal-attainment failures. We operationalized failure as the experience of pursuing but not succeeding at a focal task. The failure tasks we used had participants working toward difficult goals, often with high standards for performance and performance implications (e.g., successful workplace and interpersonal interactions). In the crucial conditions of the current work, we therefore induced effort—but effort that led to the experience of not succeeding at reaching the goal.

We reasoned that the realistic appraisal process associated with self-affirmation would, when facing setbacks, alter self-perceptions of efficacy in a downward fashion, such that people would doubt their goal-attainment abilities. The perception that one is unable to achieve a goal was proposed to lead to (and account for) poor goal performance. In contrast, when people perform tasks that are not associated with failure, self-affirmation was expected to bolster a sense of self-efficacy and improve performance.

We tested our hypotheses in four experiments. Experiment 1 manipulated self-affirmation and then had participants think about the self’s abilities to pursue goals as a key reason why self-affirmation leads to goal disengagement in the context of failure.

Experiment 1: Readiness to Disengage

Experiment 1 tested whether being self-affirmed alters tendencies to disengage from current life goals. We asked affirmed and nonaffirmed participants to think of a current goal in their lives that was important to them. Then they were asked to simulate what it would be like to have to stop pursuing that goal. Simulation often is psychologically very similar to responding to the situation behaviorally (Baumeister, Masicampo, & Vohs, 2011), and so we used that method to mimic the reality of halting goal pursuit. Responses on the disengagement subscale from the Goal Adjustment Scale (Wrosch, Scheier, Miller, Schulz, & Carver, 2003) were our primary dependent measure.

This initial test of our hypothesis also included trait measures that prior work has associated with goal disengagement so that we could assess the effect of self-affirmation above and beyond the contributions of these traits. Self-esteem, dispositional optimism, and approach and avoidance tendencies have been linked to differences in goal-dissatisfaction tendencies (Aspinwall & Richter, 1999; Bauer & Wrosch, 2011; Wrosch & Miller, 2009) and so we included them in Experiment 1. We also included a state self-awareness scale to test whether the self-affirmation manipulation piques momentary self-awareness, which can promote self-regulation (Carver & Scheier, 1990) and therefore might heighten goal-dissatisfaction tendencies.

We predicted that self-affirmation would have a significant, positive effect on reports of goal-dissatisfaction responses relative to a no-affirmation condition. We expected that the self-affirmation effect would be robust when examined in the context of traits associated with goal disengagement. We also did not expect changes in state self-awareness, as prior research has found that self-affirmation may alter higher order changes in self-perception (e.g., feelings of vulnerability; Klein et al., 2011) but does not necessarily induce excessive awareness of the self.

Method

Participants. Fifty-two adults (19 female; M age = 31.54 years) participated in exchange for partial course credit. They were randomly assigned to condition in a 2 (self-affirmation vs. no affirmation) × 2 (affirmation first vs. goal selection first) between-subjects design. The latter factor involved manipulating whether participants nominated an important goal in their lives before or after the self-affirmation task. This order factor was used to assess whether engaging in self-affirmation would alter the nature of the goals that participants nominated in terms of whether they were easier or harder from which to disengage. As expected, order had no effect (main effects and interaction Fs < 1), and so we do not discuss this factor.

Procedure. Participants completed the experiment in a group setting. They first completed several trait measures that have been related to goal disengagement. Global self-esteem was measured with Rosenberg’s (1965) popular 10-item questionnaire, with 0 = strongly disagree, 3 = strongly agree; α = .81. Dispositional optimism was measured with Scheier, Carver, and Bridges’s (1994) Life Orientation Test–Revised, for which participants used a 5-point scale (1 = strongly disagree, 5 = strongly agree; α = .78). Approach and avoidance tendencies were measured with 2 six-item scales from Elliot and Thrash (2002), on which participants rated their agreement using 7-point scales (1 = strongly
Participants next considered nine personal values and characteristics that people may deem important. The list included athletics, business and managerial skills, and relationships with friends and family (borrowed from Cohen, Aronson, & Steel, 2000). As is customary in self-affirmation research, all participants ranked the values. Then participants in the self-affirmation condition wrote several lines explaining why their top-ranked value was important to them and a time in their lives when it had been particularly important. Participants in the no-affirmation condition wrote several lines describing why and when the value they had ranked fifth (i.e., middling) in importance played a role in their lives (Fein & Spencer, 1997; Sherman et al., 2000). After the values affirmation task, participants completed the Situational Self-Awareness Scale (Govern & Marsch, 2001), which has nine items to which participants respond using a 7-point scale (1 = strongly disagree, 7 = strongly agree; \( \alpha = .69 \)).

Last, participants completed the dependent measure, the disengagement subscale from the Goal Adjustment Scale (Wrosch et al., 2003). Participants described a personal goal, such as trying to lose weight or score well on upcoming exams, that was important to them and that they were currently attempting to attain. Participants’ nomination of personal current goals was different from the writing task done earlier, which was to reflect on a value chosen from a list of preestablished values (Cohen et al., 2000; Fein & Spencer, 1997; Sherman et al., 2000).

Next, participants read and followed the instructions from the Goal Adjustment Scale, which called for them to imagine how they would respond if they would have to stop pursuing their personal goal (four items; \( \alpha = .65 \)). Items included “It would be easy for me to reduce my effort toward the goal” and “I would stay committed to the goal for a long time; I wouldn’t be able to let it go” (reverse coded). Participants rated their response to each item using a scale from 1 to 5 (definitely not to definitely so). After completing the Goal Adjustment Scale, participants were debriefed.

Results and Discussion

We first tested the hypothesis that self-affirmation predicts goal disengagement. A \( t \) test showed that it did, \( t(50) = 2.26, p < .03 \). Participants assigned to the self-affirmation condition (\( M = 3.17, SD = 0.49 \)) reported that they would more readily disengage from a life goal compared to participants in the nonaffirmation condition (\( M = 2.85, SD = 0.52 \)).

We next tested whether the effect of self-affirmation condition predicted goal engagement above and beyond the influence of other variables that have been tied to goal regulation. Separate regression analyses predicted scores on the goal-disengagement scale from self-affirmation condition alongside measures of avoidance, approach, optimism, and self-esteem, respectively. In all cases, the effect of self-affirmation remained significant (\( ts > 2.02, ps < .05 \)). None of the personality traits was significantly associated with goal disengagement (\( ts < 1.5 \)), nor did any of the personality variables interact with self-affirmation condition (\( ts < 1 \)). These patterns suggest that the effect of self-affirmation on goal disengagement was not redundant with other plausible goal adjustment processes.

Next, we conducted a \( t \) test to assess whether self-affirmation enhances goal disengagement by increasing self-awareness. As predicted, self-affirmation condition was not a significant predictor of scores on the Situational Self-Awareness Scale (\( t < 1 \)), suggesting that self-affirmation’s effect on disengagement was not due to an uptick in self-awareness.

In summary, engaging in self-affirmation made participants feel that they would have an easier time walking away from a current and important goal in their lives. The effect was robust to order effects in procedure, remained significant even when other goal adjustment variables were included in the analyses, and was not due to momentary increases in self-awareness.

While offering supportive results, Experiment 1, however, relied on a hypothetical scenario to gauge disengagement rather than testing behavior in a real performance setting (cf. Baumeister, Vohs, & Funder, 2007). Accordingly, the next experiment assessed actual performance in a failure versus no-failure situation.

Experiment 2: The Importance of Expectations

Experiment 2 tested the hypothesis that participants who had been affirmed on their core values would perform poorly when attempting a task riddled with failure. After a self-affirmation manipulation, participants were instructed to move individual pieces of rice with chopsticks. In the failure version of the chopsticks task, participants were instructed to perform quickly and well. Because the vast majority of our participants had little experience with chopsticks, we expected that this version of the task would be beset with failure. (We were correct in this assumption.) In the no-failure version of the task, participants were expected simply to give the chopsticks task a try, without performance expectations. After an initial attempt at the task induced the experience (or nonexperience) of failure, we measured three outcomes: performance expectations for subsequent attempts at the task, actual performance on subsequent attempts, and motivation to continue with a similar task.

We predicted that when the chopsticks exercise was devoid of failure, being self-affirmed would benefit motivation and performance. This prediction follows from research showing that self-affirmation can enhance the desire to reach one’s goals and produce successful behavioral change when the goal is relatively straightforward to achieve. For instance, self-affirmation has proven successful in getting people to eat more fruits and vegetables than they would otherwise (Epton & Harris, 2008). Therefore, we anticipated that the self-affirmation manipulation would improve motivation on and performance of a task that was not associated with failure.

Unique to the current research were predictions about when self-affirmation would produce poor motivation and performance. We predicted that affirming the self would reduce performance expectations and actual performance on the task when initial attempts were met with failure. Moreover, we predicted that affirming the self would cause participants to withdraw effort from further attempts at the task, consistent with the prediction that self-affirmation can hasten goal disengagement.

Method

Participants. One hundred thirty-two undergraduates (95 female; \( M \) age = 21.44 years) were paid $10 for participating.
Participants were randomly assigned to condition in a 2 (self-affirmation vs. no affirmation) × 2 (failure vs. no failure) between-participants factorial design.

Self-affirmation manipulation. Participants rank-ordered 11 values and characteristics in terms of personal importance. As in Experiment 1, participants then wrote about one of those values (Cohen et al., 2000; Fein & Spencer, 1997; Sherman et al., 2000). Participants in the self-affirmation condition wrote a brief essay explaining the importance of their top-ranked value and when in their lives it had been particularly important. Participants in the no-affirmation condition wrote about why and when the value they had ranked seventh (i.e., middling) in importance would be important to the average college student.

Failure manipulation. Next participants performed a task that involved moving individual pieces of rice using chopsticks. This task is difficult for people who do not use chopsticks very often. A pretest (n = 30) conducted with participants from the same population as in the main experiment verified the difficulty of this task. Participants in the pretest were given an opportunity to move 20 pieces of rice using chopsticks, one at a time, between two places that were placed 25 cm apart. After 90 s, the average number of rice pieces participants had moved was fewer than three (M = 2.43, SD = 2.36). This result confirmed that moving individual pieces of rice with chopsticks would result in failure for the vast majority of our participants.

In the main experiment, participants in the failure condition were told that the task was a predictor of thinking on one’s feet, a skill that is valuable in both the workplace and interpersonal interactions. Participants were instructed to use chopsticks to move 20 rice pieces, one at a time, between two plates that were placed 10 in. apart, within 90 s. None of the participants in this condition reached the goal, and hence all experienced failure. Participants in the no-failure condition were told that the task was a pretest for another study. They were instructed to “give it a try” and move 20 pieces of rice, one at a time, between two plates until the experimenter asked them to stop. These participants were not given a time limit or a performance goal, nor were there stated implications for their performance (Vohs & Heatherton, 2001). We assumed participants in this condition would not perceive that they failed (or succeeded) at the task.

Performance expectations for upcoming attempts. After the first attempt, we measured participants’ expectations for performance on upcoming attempts. They reported whether they expected to do better on the next attempts by marking an 11-cm line anchored by endpoints of not at all and very much. Higher numbers indicated loftier performance expectations.

Subsequent performance. Next, participants performed the chopsticks task twice more. The sum of the number of rice pieces that participants moved across these two attempts (α = .84) was the outcome measure of performance.

Interest in attempting a similar task. After the chopsticks task, we assessed participants’ desire to put additional effort toward a similar task. Participants learned about a device that helps people use chopsticks (which we labeled a chopsticks helper). They reported how interested they were in attempting another rice-moving task using the chopsticks helper (1 = not at all, 7 = very much).

Prior chopsticks use. We expected participants who had more experience with chopsticks to perform better than those who were unfamiliar with using chopsticks. Hence, we asked participants to report how often they used chopsticks on a scale from 1 (never) to 5 (every day; Mode = 1; M = 2.32, SD = 1.37).

Results

Manipulation check pretest. To verify that the failure condition elicited more failure-related perceptions compared to the no-failure condition, we conducted a pretest with 50 participants from the same population as participants in the main experiment. Pretest participants performed the failure versus no-failure versions of the chopsticks task after either self-affirmation or no self-affirmation. A 2 (self-affirmation) × 2 (failure) analysis of covariance (ANCOVA) with frequency of chopsticks use as a covariate compared ratings of task difficulty and degree of frustration. As expected, participants in the failure condition reported that their task was more difficult (M = 6.43, SD = 0.80) and frustrating (M = 5.58, SD = 1.36) than did participants in the no-failure condition (difficult: M = 5.22, SD = 1.67; frustrating: M = 4.71, SD = 1.45), F(1, 45) = 13.89, p < .01, and F(1, 45) = 4.47, p < .05. Prior chopsticks usage was also a significant predictor, F(1, 45) = 28.82, p < .001. The main effect of affirmation condition was not a significant predictor of difficulty or frustration ratings, nor was the Affirmation Condition × Failure Condition interaction (Fs < 1.60). These pretest results confirmed that the failure version of the chopsticks task did indeed induce greater experience of failure.

Performance expectation for upcoming task attempts. After completing an initial round of the chopsticks task (which induced failure or nonfailure), participants in the main experiment reported their performance expectation for upcoming attempts at the task. We predicted that affirming core values would reduce expectations for success on subsequent attempts at the chopsticks task when the first attempt had met with failure. Our prediction was supported by the results of a 2 (self-affirmation) × 2 (failure) ANCOVA with frequency of chopsticks use as a covariate. As predicted, the interaction of the two factors was a significant predictor of performance expectations for subsequent attempts at the chopsticks task, F(1, 127) = 9.62, p < .01, and prior chopsticks usage, F(1, 127) = 6.19, p < .05. There were no significant main effects of self-affirmation condition (F < 1) or failure condition, F(1, 127) = 2.21, p > .13.

 Planned comparisons revealed that, on the one hand, participants who affirmed a cherished value and then performed a task devoid of failure had higher performance expectations for subsequent task attempts (M = 6.21, SD = 3.00) compared to those who did not affirm a cherished value (M = 4.53, SD = 2.85), F(1, 127) = 5.38, p < .05. These results are consistent with prior self-affirmation work. On the other hand, the predicted dampening effect emerged among self-affirmed participants who performed the failure-laden version of the task (M = 4.04, SD = 2.50), as opposed to their counterparts who did not self-affirm (M = 5.34, SD = 2.78), F(1, 127) = 4.24, p < .05. From another perspective, the data showed that self-affirmed participants expected to perform less well after failure than after the no-failure task, F(1, 127) = 10.07, p < .01. Their counterparts who did not self-affirm showed no such difference F(1, 127) = 1.36, p > .24. Hence, self-affirmation’s effects on performance expectations diverged as a
function of whether participants had experienced failure (deflated expectations) or nonfailure (boosted expectations).

**Performance on the second and third attempts of the chopsticks task.** Our main hypothesis was that affirming one’s core values harms performance when facing a task imbued with failure. This hypothesis was supported by the results of a 2 (self-affirmation) × 2 (failure) ANCOVA, with chopsticks-use frequency as a covariate, predicting performance on the latter two attempts at the chopsticks task. As hypothesized, the interaction of the two factors was a significant predictor of performance (see Figure 1), $F(1, 127) = 8.62, p < .01$, as was prior chopsticks usage, $F(1, 127) = 27.91, p < .001$. There were no significant main effects of self-affirmation condition ($F < 1$) or failure condition, $F(1, 127) = 1.10, p > .29$.

Planned comparisons revealed the predicted effects. On the one hand, having affirmed a cherished value before performing the failure-free version of the chopsticks task led to better performance ($M = 12.03, SD = 13.20$) compared to those who did not affirm a cherished value ($M = 6.47, SD = 6.70$), $F(1, 127) = 4.26, p < .05$. This replicates past work. On the other hand, the dampening effect emerged among self-affirmed participants who performed the failure-laden version of the task ($M = 5.66, SD = 7.91$), as opposed to their counterparts who did not self-affirm ($M = 9.80, SD = 11.96$), $F(1, 127) = 4.34, p < .05$. From another perspective, self-affirmed participants performed worse after failure than after nonfailure, $F(1, 127) = 7.59, p < .01$. Their counterparts who did not self-affirm showed no such difference, $F(1, 127) = 1.87, p > .17$. Hence, the effects of self-affirmation diverged as a function of the experience of failure. It helped performance among those who had not experienced failure, whereas it harmed performance for those who had previously experienced failure at the task.

**Mediation analyses.** We tested whether changes in performance expectation could account for changes in actual performance on the subsequent chopsticks tasks as a function of self-affirmation condition and failure condition. The first two qualifications for mediation had already been met: The interaction of self-affirmation and failure predicted subsequent performance on the chopsticks task (the dependent measure), $b = -.99, \tau(127) = 2.94, p < .01$, and performance expectation (the proposed mediator), $b = -1.12, \tau(127) = 3.10, p < .01$. Next, we assessed whether performance expectations correlated with subsequent performance, which they did, $b = .38, \tau(129) = 5.03, p < .001$, as did prior chopsticks usage, $b = .34, \tau(129) = 4.47, p < .001$. When the interaction of self-affirmation condition and failure condition and the proposed mediator (performance expectations) were regressed on chopsticks-task performance, performance expectations remained a significant predictor, $b = .34, \tau(126) = 4.29, p < .001$, whereas the significance level of the interaction of self-affirmation and failure dropped, $b = -.62, \tau(126) = 1.88, p > .05$. A Sobel’s test confirmed significant mediation of chopsticks-task performance by performance expectations ($z = 2.53, p < .05$). In summary, changes to beliefs about the likelihood of achieving one’s goals appeared to be responsible for changes in actual performance after participants had received a self-affirmation treatment (or not) and failed at a task (or not).

**Interest in performing a similar task.** After completing the chopsticks task, participants indicated how interested they were in trying a nearly identical task (i.e., using the chopsticks helper to move individual pieces of rice). We treated participants’ responses as a measure of goal motivation. A 2 (self-affirmation) × 2 (failure) ANCOVA showed the predicted interaction effect, $F(1, 127) = 9.45, p < .01$; see Figure 2. There were no significant main effects of self-affirmation condition ($F < 1$) or failure condition, $F(1, 127) = 3.03, p > .08$. The covariate, frequency of chopsticks use, was significantly related to interest in performing the task, $F(1, 127) = 6.50, p < .05$, such that participants who had less experience with chopsticks were more interested in trying the task again.

Planned comparisons revealed a reliable effect of self-affirmation in the no-failure condition, such that participants who affirmed a cherished value showed more interest in performing an additional chopsticks task ($M = 5.77, SD = 1.31$) than those who did not affirm a cherished value ($M = 5.00, SD = 1.94$), $F(1, 127) = 3.99, p < .05$. As predicted by the current model, this effect reversed for participants in the failure condition. Having affirmed a cherished value dampened their interest in performing the chopsticks helper task ($M = 4.34, SD = 2.01$) as compared to not doing so ($M = 5.37, SD = 1.55$), $F(1, 127) = 5.51, p < .05$. Additional comparisons revealed that self-affirmed participants showed reduced interest in using the chopsticks helper when their prior experience with the chopsticks task involved failure, relative to when it did not, $F(1, 127) = 11.09, p < .01$. Participants who did not self-affirm did not show this difference, $F(1, 127) < 1$. In accordance with predictions, then, the desire to try a similar task changed with self-affirmation and failure treatments, such that self-affirmation made participants in the nonfailure condition want to try the task more than they would otherwise, whereas participants in the failure condition reported less motivation.

**Discussion**

Experiment 2 tested whether the experience of failure could reverse the beneficial effects of self-affirmation typically observed on performance and motivation. The results supported these predictions. We found that affirming core values, as opposed to not affirming core values, hurt performance expectations and actual performance on a failure-laden task involving chopsticks. In contrast, affirming one’s core values prior to attempting a nonfailure version of the task improved performance expectations and actual performance relative to participants who did not self-affirm, a finding consistent with prior work on the benefits of self-affirmation.
affirmation (Epton & Harris, 2008; Martens et al., 2006; Sherman et al., 2000).

The effects of self-affirmation and the experience of failure went beyond poor performance. We found that engaging in self-affirmation and experiencing failure sapped participants’ motivation to devote additional effort to the task. Affirmed participants who experienced failure were the least interested in using a chopsticks helper on a subsequent rice-moving task, an effect that meshes well with what people do when they disengage from a goal. Participants who were self-affirmed and who faced a non-failure version of the task, though, maintained interest in the task as evidenced by a strong desire to try the chopsticks helper.

Moreover, we found that the combination of self-affirmation and the initial experience of failure changed participants’ beliefs about their ensuing performance, in that affirmed participants expected poorer performance out of themselves on upcoming attempts at the chopsticks task relative to the other groups. Once again, that pattern stands in contrast to affirmed participants who did not experience failure, who reported the highest expectations about their upcoming performance.

In summary, the results of Experiment 2 revealed that self-affirmation can help or hurt performance. Its helpfulness came in the form of better outcomes and motivation to persist at a task that was not associated with failure. However, self-affirmation curbed expectations, performance, and motivation when people attempted a failure-laden task. Withdrawing effort, dampening performance expectations, and disinterest in performing similar tasks are hallmarks of goal disengagement (Wrosch et al., 2003), in which people withdraw effort on and desire for tasks that seem unlikely to end profitably.

**Experiment 3: The Importance of Goal Attainability**

Experiment 1 found that people who had recently thought about their most cherished values in life said that they would more easily disengage from an obstinate goal compared to participants who had not thought deeply about their most cherished value. Experiment 2 found that if participants had been self-affirmed and experienced failure during an initial performance, their expectations for future success at the task dropped, as did their motivation to attempt a similar task. Perhaps most striking was that the combination of self-affirmation and failure harmed actual performance on the task, which was explained (using mediation) by dampened performance expectations.

Experiment 3 was performed with two central aims in mind. First, whereas Experiment 2 measured motivation and performance on a task that was (for some participants) associated with failure from the start, Experiment 3 tested the generality of the effect by asking whether the deleterious effect of failure on one task could carry over to affect motivation and performance on another task. Hence the current experiment used a behavioral dependent measure that shared vague qualities with the task that manipulated the experience of failure but was different enough to offer a rigorous test of our hypotheses.

Second, we examined the role of perceived likelihood of goal attainment in the expected behavioral outcomes. Our theory is that affirmed participants accept self-relevant information and integrate it into their self-concepts. The notable case, as in much prior research, involves the receipt of negative information (such as the perception one may not attain a current goal). If self-affirmation aids in integrating self-relevant information—even that which is unfavorable—and if this informs future behavior, then perceiving failure on one goal should result in disengagement and demotivation on other goals that are believed to rely on similar capacities. Experiment 2 revealed correlational evidence consistent with this view using self-reports of performance expectations as a statistical mediator. In Experiment 3, we manipulated perceptions of the likelihood of goal attainment in the interest of drawing firmer causal conclusions.

All participants learned that a chopsticks task (the manipulation) was an indicator of thinking on one’s feet. Afterwards, participants completed another task that also was said to be an indicator of thinking on one’s feet. Depending on experimental condition, participants were given the impression that achievement on these tasks was likely or unlikely to be attained. We predicted that when success was unlikely, self-affirmed participants would show less motivation and effort compared to their nonaffirmed counterparts, whereas they would show increased motivation and effort when goal attainment was likely.

**Method**

**Participants.** One hundred nineteen undergraduates (70 female; M age = 20.18 years) participated in exchange for extra course credit. Participants were randomly assigned to condition in a 2 (self-affirmation vs. no affirmation) × 2 (attainable goal vs. unattainable goal) between-participants factorial design.

**Affirmation-condition manipulation.** As in Experiment 2, participants ranked 11 values in order of personal importance. Next, participants in the self-affirmation condition wrote a short essay about why their top-ranked value was important to them and described a time in their lives when it had been particularly important. Participants in the no-affirmation condition wrote a short essay about why and when the value they ranked seventh (i.e., middling) in importance might be important to the average college student.

**Goal-attainability manipulation.** Next participants performed two tasks purported to be good predictors of the capacity to think on one’s feet. The first was a rice-moving task similar to the one used in Experiment 2, and the second was a remote associates test (RAT; Mednick, 1968), which is described in detail below. The tasks and instructions participants received varied as a function of experimental condition.
Participants in the attainable goal condition were given the moderately challenging goal of using chopsticks to move 10 rice pieces, one at a time, between two plates that were 10 in. apart, all within 90 s. They were told that this task was very challenging but probably they would succeed at it if they tried hard enough. We reasoned that, like the no-failure version of the chopsticks task used in Experiment 2, participants in this condition should experience relatively little failure.

After the chopsticks task, participants attempted an RAT that required finding one word that linked three seemingly unrelated words. The RAT can be made relatively easy or quite difficult depending on the words used. For example, an easy item would be shopping, washer, and picture, all of which are associated with window. Participants in the attainable goal condition attempted to solve 12 RAT items, of which six were relatively easy and six were moderately difficult (Kihlstrom, Shames, & Dorfman, 1996, validated the items’ difficulty). In keeping with the goal-attainability manipulation, participants in the attainable goal condition again were told that the task was very challenging but they probably would succeed if they tried hard enough.

Participants in the unattainable goal condition were given the demanding goal of using chopsticks to move 20 rice pieces, one at a time, between two plates that were 10 in. apart, all within 90 s. Furthermore, they were told that they would probably feel unsure of whether they could succeed. We reasoned that, like the failure version of the chopsticks task used in Experiment 2, participants in this condition would consistently experience failure. After the chopsticks task, participants attempted to solve a set of 12 highly difficult RAT items (from Kihlstrom et al., 1996). In keeping with the goal-attainability manipulation, participants in the unattainable goal condition again were told that the task was very challenging and they would probably feel unsure of whether they could succeed.

Goal-achievement perceptions. After performing the chopsticks task but before performing the RAT, participants rated the extent to which they thought they had achieved their goal on the chopsticks task (“During the task . . . how able were you to reach the goal?/did you have a sense of accomplishment/?do you think that you reached the goal?”; α = .95) using a 10-cm line anchored with not at all and all the way. This was a manipulation check of whether participants deemed the goal as being attainable.

Expended effort on the task. Next, participants reported how much effort they expended on the chopsticks task using a scale from 1 (none) to 7 (a lot).

Intended effort on another attempt of the chopsticks task. As a measure of motivation, we assessed participants’ predictions of how much effort they would expend right then on another attempt at the chopsticks task on a scale from 1 (none) to 7 (a lot).

Prior chopsticks use. We asked participants to report how often they use chopsticks (1 = never, 5 = everyday; M = 2.13, SD = 1.16; Mode = 1). We expected that participants with more experience using chopsticks would find the first task less difficult than participants unfamiliar with using chopsticks, and therefore we controlled for prior chopsticks use.

After performing the chopsticks task, responding to the above questions, and attempting their respective RAT items, participants were debriefed, thanked, and dismissed.

Pretest 1: Thinking on one’s feet. We conducted a pretest with 61 participants from the same population as participants in the main experiment to ensure that the chopsticks task and RAT both could be viewed as relying on the same skill (i.e., thinking on one’s feet). Pretest participants first completed the self-affirmation versus no-affirmation induction. They next read about the chopsticks task and RAT and, depending on condition, were given the same (attainable vs. unattainable) information about chances of success as in the main study. They reported how much each task reflected the ability to think on one’s feet, 1 (not at all) to 7 (very much).

As expected, both tasks were perceived as good reflections of thinking on one’s feet (chopsticks task: M = 3.36, SD = 1.49; RAT: M = 4.62, SD = 1.60), with both ratings significantly different from 1 (not at all): chopsticks task, t(60) = 12.33, p < .001; RAT, t(60) = 17.64, p < .001. A 2 (self-affirmation vs. no affirmation) × 2 (goal attainable vs. unattainable) ANCOVA with frequency of chopsticks use as a covariate revealed that participants’ perceptions of the tasks were unaffected by either main effect or the interaction of the two factors (ps > .18). Both the chopsticks task and RAT therefore could be viewed as tasks that called upon the same skill set, namely, thinking on one’s feet, independent of self-affirmation or goal-attainability manipulations.

Pretest 2: Attempted answers signal motivation. Because we gave participants different RAT items by condition (easier for the attainable goal condition and more difficult for the unattainable goal condition), we could not directly compare performance across conditions. In addition, the highly difficult RAT items were so demanding for our target population that they had very low scores and little variance in number of problems correctly solved. Accordingly, we used the number of RAT items participants attempted to solve as a measure of goal-related effort.1

Results

Manipulation check. In the main experiment, we first checked whether the goal-attainability manipulation altered perceptions of goal achievement on the chopsticks task. As expected, a one-way ANCOVA with frequency of chopsticks use as a covariate revealed a significant main effect of goal-attainability condition, F(1, 116) = 23.91, p < .001 (attainable goal condition

1 To verify that the number of attempted answers on the RAT was a good indicator of motivation, we conducted a pretest with 38 participants from the same population as in the main experiment. Pretest participants were shown the actual test results of two students. Student A attempted nine questions out of 12 questions, three of which were correct answers. Student B attempted only four questions, three of which were correct answers. Participants rated which student was more motivated and which tried hard to do well (1 = Student A, 4 = similar, 7 = Student B). We also assessed participants’ perception of the students’ performance by asking which student performed objectively better (1 = Student A, 4 = similar, 7 = Student B). We predicted that compared to Student B, Student A would be viewed as being more motivated to do well on the RAT task even if both students’ total numbers of correct responses were objectively the same. As expected, participants perceived the student with higher number of attempted answers on the RAT (Student A) as the more motivated student. Their average rating (M = 2.42, SD = 1.62) was significantly below the scale midpoint of 4, t(37) = 6.00, p < .001. Also, they reported that Student A tried harder to do well on the RAT task than Student B (M = 2.53, SD = 1.98), which too was significantly different from the scale midpoint of 4, t(37) = 4.58, p < .001. Note that the performances of these two students were objectively viewed as similar (M = 3.76, SD = 1.85). Hence, pretest participants believed that of two students with an equal number of correct answers, the student who attempted more answers was indeed more motivated on the task.
M = 4.95, SD = 3.84; unattainable goal condition M = 2.26, SD = 2.20). The covariate, chopsticks use, was also a significant predictor of perceived goal achievement, F(1, 116) = 10.47, p < .01, with more frequent use relating to greater goal-achievement perceptions. Also, a 2 (self-affirmation vs. no affirmation) × 2 (goal attainable vs. unattainable) ANCOVA showed that neither the main effect of self-affirmation condition nor the Affirmation × Goal Attainability interaction was significant (Fs < 1). Therefore, as intended, the goal-attainability manipulation altered participants’ feelings of success at the chopsticks task.

**Effort expended on the chopsticks task.** We tested whether engaging in self-affirmation prior to attempting the chopsticks task altered the effort participants put forth on the task as a function of goal condition. For each goal condition, we conducted a separate one-way ANCOVA with chopsticks-use frequency as a covariate.

In the attainable goal condition, the results of a one-way ANCOVA confirmed the beneficial effects of self-affirmation. Having affirmed a cherished value before attempting a task that was seen as attainable led participants to expend more effort (M = 6.20, SD = 0.76) compared to their counterparts who did not affirm a cherished value (M = 5.50, SD = 1.31), F(1, 52) = 7.77, p < .01, as was prior chopsticks use, F(1, 52) = 4.35, p < .05.

The results within the unattainable goal condition also supported our predictions. Self-affirmation had a dampening effect on effort among participants who attempted a task in which the goal was seen as likely unattainable (M = 4.97, SD = 1.79), as opposed to participants who had not been self-affirmed (M = 5.65, SD = 1.20), F(1, 61) = 4.02, p < .05. Prior chopsticks use was not significantly related to the effort expended on the chopstick task, F(1, 61) = 2.17, p > .14.

**Intended effort on another attempt at the chopsticks task.** As an indicator of motivation, we assessed how much effort participants said they would exert on another chopsticks task (mediator), which then enhanced goal engagement on a different but related task. We next tested whether changes in motivation to continue pursuing the initial goal (i.e., effort intended to perform another chopsticks task) could account for changes in goal motivation on a new task (i.e., number of RAT problems attempted) that was said to reflect the same skill (i.e., thinking on one’s feet). We tested this prediction again with separate analyses for each goal-attainability condition.

In the **goal-attainability condition**, self-affirmation condition (no affirmation = 0, self-affirmation = 1) predicted the number of RAT problems attempted (the key dependent variable), b = .30, t(52) = 2.20, p < .05. Self-affirmation also predicted the effort participants said they would exert on another chopsticks task (proposed mediator), b = .42, t(52) = 3.28, p < .01. Next we assessed whether the effort participants said they would exert correlated with number of RAT problems attempted; it did, b = .57, t(52) = 5.00, p < .001, but prior chopsticks usage did not, b = .01, t(52) < 1.

When self-affirmation and intended effort were regressed on the number of RAT problems attempted, the effect of the mediator (intended effort) remained significant, b = .55, t(51) = 4.30, p < .001, while self-affirmation condition fell to nonsignificance, b = .07, t(51) < 1, Sobel z = 2.60, p < .01. When goals were attainable, affirming core values enhanced motivation on the initial task, which then enhanced goal engagement on a different but related task.

In the **goal-unattainability condition**, self-affirmation condition (no affirmation = 0, self-affirmation = 1) predicted the number of RAT problems attempted (dependent variable), b = −.25, t(61) = 1.96, p = .055. Self-affirmation also predicted the effort participants said they would exert on another chopsticks task (mediator), b = −.29, t(61) = 2.44, p < .05. Next we assessed whether the effort participants said they would exert correlated with number of RAT problems attempted; it did, b = .47, t(61) = 4.02, p < .001, but prior chopsticks usage did not, b = −.01, t(61) < 1.

When self-affirmation condition and intended effort were regressed on the number of RAT problems attempted, the effect of the mediator (intended effort) remained significant, b = .44, t(60) = 3.54, p < .01, while self-affirmation condition fell to nonsignificance, b = −.12, t(60) < 1, p > .33, Sobel z = 2.01, p < .05. When goals were potentially unattainable, affirming core values deflated the motivation to pursue the initial goal, which diminished goal engagement on a different but related task.

**Number of RAT items attempted to solve.** We predicted that self-affirmation, combined with goal condition, would alter participants’ efforts toward achieving a new but related goal, namely, attempting to solve RAT items. We predicted that having engaged in self-affirmation would heighten it in the attainable goal condition while lowering effort in the unattainable goal condition.

In the unattainable goal condition, the expected effects were supported by a one-way ANCOVA with chopsticks-use frequency as a covariate.\(^2\) Compared to those who did not affirm the self, those who were self-affirmed exerted less effort (i.e., attempted fewer RAT problems) when their initial task was framed as unattainable (M = 10.06, SD = 3.48, vs. M = 8.30, SD = 4.55), F(1, 61) = 3.82, p = .055.

The same ANCOVA in the attainable goal condition revealed the predicted reversal. These results showed that having affirmed a cherished value led to enhanced effort, in that these participants attempted more RAT problems (M = 10.92, SD = 1.61) than did participants who had not affirmed a cherished value (M = 9.10, SD = 4.23), F(1, 52) = 4.86, p < .05.

**Mediation analyses.** We next tested whether changes in motivation to continue pursuing the initial goal (i.e., effort intended to perform another chopsticks task) could account for changes in goal motivation on a new task (i.e., number of RAT problems attempted) that was said to reflect the same skill (i.e., thinking on one’s feet). We tested this prediction again with separate analyses for each goal-attainability condition.

\(^2\) We consistently used frequency of chopsticks use as a covariate because it affects perceived difficulty for the initial task, which could carry through to affect subsequent tasks.
Discussion

Self-affirmed participants responded differently when they had been induced to view a goal as attainable versus potentially unattainable. In the former case, being self-affirmed (vs. not affirmed) led to enhanced motivation and persistence. In the latter, being self-affirmed (vs. not affirmed) warped motivation and lowered persistence. Participants in the unattainable goal condition felt that they were unsuccessful at achieving their initial goal involving chopsticks and subsequently reported disinterest in performing the task again. This deflated motivation causally affected (demonstrated via mediation) their performance on a new task involving word puzzles, insofar as they attempted fewer items.

In all, the effect of self-affirmation on motivation and performance seems divisive in that it helps goal engagement when tasks are seen as attainable and are not associated with failure, while also diminishing motivation, harming performance, and infecting performance on other tasks when goal completion is questionable and attempts are met with failure.

Experiment 4: The Importance of Self-Efficacy

Experiment 4 tested again why affirming the self can harm performance and motivation on tasks beset with failure. We posited that because self-affirmation induces an honest assessment of the circumstance and one’s role in the experience of failure (Kihlstrom et al., 1996), Participants in this condition were told that the test is commonly used to assess intelligence and that they had 8 min to complete the items. Participants in the no-failure condition attempted easy-to-solve RAT items and also worked for 8 min. For these participants, the idea of assessing intelligence was not mentioned nor was the time limit. Similar instructions have been used in past research to manipulate the experience of failure (Heatherton & Vohs, 2000; Vohs & Heatherton, 2001). As a manipulation check, participants rated how difficult and frustrating the RAT was using scales from 1 (not at all) to 7 (very much).

Self-perception measure. Participants next completed a 25-item measure of self-perceptions. This questionnaire listed two opposing traits as endpoints of a 7-point scale. The traits were anxious/calm, arrogant/modest, bold/timid, cold/warm, conscientious/careless, cruel/kind, energetic/lethargic, friendly/unfriendly, gloomy/cheerful, happy/sad, helpless/helpful, high self-esteem/low self-esteem, deceptive/dishonest, incapable/capable, incompetent/competent, inconsiderate/considerate, intelligent/unintelligent, irresponsible/responsible, lazy/hardworking, polite/rude, positive/negative, sincere/fake, sophisticated/crude, thoughtful/thoughtless, and unattractive/attractive. Participants were instructed to think of how they felt about themselves at that very moment.

Subsequent performance. Participants’ final task was to complete a new set of RAT items as a measure of performance following the experience or nonexperience of failure. Participants spent 8 min attempting eight RAT items that were middling in difficulty. Performance was measured in terms of number of items correct.

Results

Manipulation checks. To ensure that participants in the no-affirmation condition could complete their version of the affirmation task, we asked all participants whether they were familiar with Bill Gates; all of them were. To check whether the procedures successfully elicited failure-related perceptions, an analysis of variance (ANOVA) with self-affirmation and failure conditions was conducted on ratings of task difficulty and degree of frustration. Participants in the failure condition reported that their task was more difficult (M = 4.55, SD = 2.10) and frustrating (M = 5.12, SD = 1.19) than did participants in the no-failure condition (difficult: M = 2.48, SD = 1.35; frustrating: M = 2.98, SD = 1.44), F(1, 52) = 61.52, p < .01, and F(1, 52) = 25.52, p < .01. The main effect of affirmation condition was not a significant
predictor of either rating, nor was the Affirmation Condition × Task Condition interaction ($F$s < 2.60).

**Self-perceptions.** We predicted that self-affirmation combined with the failure version of the RAT would reduce participants’ ratings of their own intelligence (which is the trait around which the manipulation was framed) and their self-efficacy perceptions generally.

Self-perceptions of intelligence were subjected to an ANOVA with self-affirmation condition and failure condition as predictors. We found no effect of affirmation condition ($F < 1$) and a marginal effect of failure condition, $F(1, 52) = 2.51, p < .12$, such that participants who received the failure version of the initial RAT saw themselves as less intelligent ($M = 3.30, SD = 1.75$) than participants who received the no-failure version ($M = 4.00, SD = 1.78$). This analysis also revealed the predicted interaction between affirmation condition and failure condition, $F(1, 52) = 12.84, p < .01$.

Planned comparisons confirmed that for participants who performed the failure version of the initial RAT, being self-affirmed led to more negative self-perceptions of intelligence ($M = 2.57, SD = 1.34$) than not being self-affirmed ($M = 4.07, SD = 1.85$), $t(52) = 2.43, p < .02$. In contrast, when participants attempted the no-failure version of the initial RAT, being self-affirmed ($M = 4.80, SD = 1.78$) was helpful in boosting self-perceptions of intelligence relative to not being affirmed ($M = 3.21, SD = 1.42$), $t(52) = 2.65, p < .05$. Another perspective is to compare ratings within self-affirmation versus no-affirmation conditions. Among participants who were self-affirmed, performing the failure versus no-failure version of the initial RAT significantly reduced self-perceptions of intelligence, $t(52) = 3.72, p < .01$, whereas, for their counterparts who were not self-affirmed, self-perceptions of intelligence were unchanged by the experience of failure ($t < 1.40, p > .17$).

We next tested whether being self-affirmed and having attempted a failure-laden task would negatively affect self-perceptions on traits related to goal pursuit, namely, ratings on the traits capable, competent, responsible, and hardworking (composite $\alpha = .64$). An ANOVA with self-affirmation condition and failure condition revealed a main effect of failure condition, $F(1, 52) = 5.42, p < .03$, with participants in the failure condition reporting lower self-efficacy as compared to those in the no-failure condition ($M = 17.63, SD = 4.67$, vs. $M = 20.17, SD = 4.02$), and no main effect of self-affirmation condition ($F < 1$). As predicted, the interaction between affirmation condition and task condition was significant, $F(1, 52) = 17.05, p < .01$; see Figure 3.

Planned comparisons indicated that within the failure condition, self-affirmation reduced self-efficacy perceptions compared to not being self-affirmed ($M = 15.42, SD = 4.26$, vs. $M = 20.00, SD = 3.98$), $t(52) = 3.09, p < .01$. Within the no-failure condition, we observed that being self-affirmed elevated self-efficacy perceptions relative to not being affirmed ($M = 22.07, SD = 2.63$, vs. $M = 18.14, SD = 4.34$), $t(52) = 2.75, p < .01$. Comparing within self-affirmation versus no-affirmation conditions, we found that for participants in the self-affirmation condition, self-efficacy perceptions were more negative after the failure versus no-failure version of the initial RAT, $t(52) = 4.65, p < .01$. There was no difference among no-affirmation participants, $t(52) = 1.26, p > .21$.

**Performance on the second set of RAT items.** We predicted that self-affirmation prior to the failure version of the initial RAT would reduce performance on a subsequent set of RAT items. A 2 (self-affirmation condition) × 2 (failure condition) ANOVA revealed the predicted interaction effect, $F(1, 52) = 6.02, p < .05$; see Figure 4. Additionally, there was a main effect of failure condition, $F(1, 52) = 15.85, p < .01$, with no-failure participants scoring higher than failure condition participants ($M = 4.38, SD = 1.68$, vs. $M = 2.74, SD = 1.43$), but no effect of self-affirmation condition ($F < 1$).

Planned comparisons confirmed that within the failure condition, self-affirmation reduced performance relative to no affirmation ($M = 2.14, SD = 1.46$, vs. $M = 3.38, SD = 1.12$), $t(52) = 2.14, p < .05$. Within the no-failure condition, in contrast, being in the self-affirmation condition tended to increase RAT performance more than not being in that condition ($M = 4.73, SD = 1.75$, vs. $M = 4.00, SD = 1.57$), $t(52) = 1.31, p < .10$ (one tailed). Analyses of RAT scores in the self-affirmation condition revealed worse performance among participants who performed the failure versus no-failure version of the task, $t(52) = 4.63, p < .01$. In the no-affirmation condition, RAT scores did not differ as a function of failure condition ($t < 1.10$).

**Mediation analyses.** We tested whether changes in self-efficacy could account for changes in performance on the second RAT as a function of self-affirmation and failure conditions. The first two qualifications for mediation already had been met: the interaction of self-affirmation and failure predicted both self-efficacy perceptions (proposed mediator) and subsequent RAT performance (dependent measure). Next, we assessed whether self-efficacy perceptions correlated with RAT scores; they did, $r(56) = .58, p < .01$. Last, we used regression models to predict RAT scores from self-affirmation condition (centered), failure condition (centered), self-efficacy perceptions (centered), and the Self-Affirmation × Failure Condition interaction. This model revealed that the putative mediator, self-efficacy perceptions, predicted RAT scores, $b = .45, r(51) = 3.60, p < .05$, and that the interaction of self-affirmation condition and failure condition was no longer significant, $b = -.07, r(51) < 1$. The main effect of failure condition was significant, $r(52) = 3.13, p < .01$, whereas

---

3 We expected conscientiousness to group with this set of traits, but reliability and factor analyses revealed that it did not cohere with the other goal-directed traits. Therefore, this trait was omitted from the efficacy index.
Discussion

The results of Experiment 4 supported the hypothesis that being self-affirmed in combination with a failure-laden task harms later performance. Participants focused on either why they held dear their top-ranked value or why Bill Gates might value neatness and tidiness. Next, participants attempted a verbal task involving word patterns that varied in whether it induced the experience of failure. Participants rated their self-concept and then attempted a second, moderately challenging version of the word puzzle.

The results revealed that self-affirmation enhanced performance of a task that did not induce failure. Participants rated their self-efficacy as being higher and performed better on a subsequent version of the task if they earlier had reflected on their personal values and completed a relatively easy version of the task. Hence, in concert with past research, we saw a beneficial effect of self-affirmation on performance and self-efficacy (Martens et al., 2006; Sherman et al., 2000; Zhao & Nan, 2010).

Nonetheless, and in line with the current theory, self-affirmation was not a uniform advantage. Focusing on one’s core values prior to attempting a failure-laden task caused participants to evaluate themselves relatively unfavorably in terms of intelligence and traits related to goal pursuit (i.e., capable, competent, energetic, and hardworking). They also performed worse than other participants on a second version of the task, an effect that was statistically accounted for by changes in self-efficacy. Therefore, because participants felt they had lower ability to achieve their goals, their performance sank when they attempted a subsequent version of the task. In total, these results suggest that what self-affirmation does to affect goal-directed behavior, it does through changes to the working self-concept.

General Discussion

The present research revealed that compared to not affirming one’s values, affirming what one values in life led to less favorable self-evaluations, deflated motivation, poorer performance, and a readiness to disengage from goals when the affirmation occurred prior to an experience of failure. In contrast, self-affirmation boosted motivation and performance on nonfailure tasks, findings that accord well with prior work on the benefits of self-affirmation (e.g., Epton & Harris, 2008; Sherman et al., 2000).

We replicated the divergent effects of self-affirmation across diverse settings, such as mentally simulating having to give up on an active goal (Experiment 1), mechanical dexterity skills (Experiments 2–3), and word puzzles (Experiments 3–4). A number of mediators for the effects were observed, including performance expectations, intended effort on upcoming tasks, and self-efficacy perceptions. These mediational effects converge to suggest that changes in beliefs about one’s abilities underlie the observed changes in performance.

The four experiments in this investigation used several different designs, tested the hypotheses on student and adult subject populations, and measured objective behavioral outcomes as well as subjective inner states. The focus on behavioral outcomes in three of the four experiments is particularly notable, as the study of actual behavior too often is neglected in psychology (Baumeister et al., 2007). This multimethod approach instills confidence in the robustness and generalizability of the conclusion that self-affirmation can help or hinder goal attainment.

Contributions to Self-Affirmation Theory

Research on self-affirmation has flourished in the past decade, with numerous studies demonstrating its capacity to change the way that people view the self, others, and circumstances. That self-affirmation reduces defensiveness and enables people to receive and heed negative information about the self has been well established (Cohen et al., 2000; Harris et al., 2007; Klein & Harris, 2009; Schmeichel & Martens, 2005; Sherman et al., 2000). The current findings demonstrate that self-affirmed people internalize the implications of failure, which in turn has major consequences for behavior.

The current results offer new perspectives on the consequences of such internationalization for motivation and performance. The present studies found that being self-affirmed led people to adopt less positive views of their goal-achievement abilities (i.e., self-efficaciousness; Experiment 4) in response to failure. We also found that affirmed persons who failed had lower expectations for future performance (Experiment 2), were less interested in performing a similar task (Experiment 2), and intended to exert less effort on an upcoming task (Experiment 3).

It is instructive to consider that nonaffirmed participants who experienced failure had more favorable self-perceptions than their self-affirmed counterparts (Experiment 4). Given that people often tend to resist the negative implications of failure, it seems that self-affirmation reduces this resistance. A relatively pessimistic sense of goal-attainment abilities among self-affirmed persons led, in turn, to a downturn in performance. The current work thus contributes to self-affirmation theory by revealing novel consequences of reduced defensiveness: When affirmed persons internalize the implications of their failed goal pursuits, it alters not only how they perceive themselves but also their behavior and priorities.

Another contribution of this work pertains to the effects of self-affirmation on health behavior. Goal disengagement may explain why self-affirmation does not reliably produce the positive changes in behavior that self-affirmed persons declare they will make. For instance, engaging in self-affirmation before reading threatening health-risk information has been shown to produce stronger intentions to quit smoking cigarettes among heavy smok-
ers, but behavioral change was absent even 1 week after the affirming event (Harris et al., 2007). Another investigation found analogous evidence among self-affirmed alcohol drinkers in the form of stated intentions to change but no real change in drinking behavior (Harris & Napper, 2005). Likewise, heavy caffeine users who read about caffeine-related health concerns showed no significant effect of self-affirmation on caffeine intake (Reed & Aspinwall, 1998).

Perhaps the gap between intention and behavior in the aforementioned contexts should not come as a surprise. It is easy to imagine that for participants who smoke, imbibe alcohol, or drink coffee regularly, curbing consumption would be beset with consistent setbacks and failures. Given that the current research found that affirmed individuals respond to failure by disengaging from the goal, we suggest that affirmation may be most effective for behavior change when initial efforts at change are not met with immediate failure. Additional research on the role of goal disengagement in understanding the disjunction between lofty intentions and a failure to enact genuine intentions seems warranted.

Limitations and Future Directions

The current experiments yielded consistent evidence for relatively more negative self-perceptions and poorer performance among affirmed persons who fail, but the boundaries of this effect have yet to be established. Do affirmed persons who fail go on to withdraw effort and disengage from any goal? Or do they disengage only from goal pursuits that result in failure, as observed in the current investigation? We found that affirmed persons who fail come to see themselves as less able to achieve their goals (Experiment 4) and that a spillover of failure perceptions to other performance contexts can occur (Experiment 3). Nonetheless, we do not think the spillover would affect performance across any and all contexts. In particular, we believe that for spillover to take place, the task should be both valuable and related to the context in which the failure occurred (Experiment 3). Additional research is needed to test this assumption.

The current experiments also did not address a crucial component of goal behavior, namely, investment in other goals. Following from discussions in the literature (e.g., Wrosch et al., 2003), we conceptualize optimal forms of goal disengagement as involving not only the abandonment of unattainable goals but also the adoption of other, more feasible goals. Rather than a net reduction in goal pursuit, affirmed persons who disengage from a failed goal might be found putting their time and energy toward new pursuits. In contrast, people who reflected on cherished values and attempted a task that was relatively free of failure outperformed participants in all other conditions. Whether people encountered failure influenced not only performance but motivation and self-perceptions, too.

Self-affirmation is an easy, straightforward act that almost anyone can do. It has a wealth of positive effects on well-being and goal pursuit. The current work demonstrates that there are circumstances in which self-affirmation’s effects on task performance and motivation are not immediately beneficial to the task at hand. Indeed, we observed that self-affirmation aids letting go of recalcitrant goals. In summary, prior research revealed that people receive a host of benefits after being self-affirmed, and the current research has uncovered a novel and nonintuitive dividend of self-affirmation in the form of goal disengagement.

References


Received February 14, 2011
Revision received August 29, 2012
Accepted August 30, 2012