Affirmed Yet Unaware: Exploring the Role of Awareness in the Process of Self-Affirmation

David K. Sherman
University of California, Santa Barbara

Geoffrey L. Cohen
University of Colorado, Boulder

Leif D. Nelson
University of California, Berkeley

A. David Nussbaum
University of Waterloo

Debra P. Bunyan
University of California, Santa Barbara

Julio Garcia
University of Colorado, Boulder

Three studies investigated whether self-affirmation can proceed without awareness, whether people are aware of the influence of experimental self-affirmations, and whether such awareness facilitates or undermines the self-affirmation process. The authors found that self-affirmation effects could proceed without awareness, as implicit self-affirming primes (utilizing sentence-unscrambling procedures) produced standard self-affirmation effects (Studies 1 and 3). People were generally unaware of self-affirmation's influence, and self-reported awareness was associated with decreased impact of the affirmation (Studies 1 and 2). Finally, affirmation effects were attenuated when people learned that self-affirmation was designed to boost self-esteem (Study 2) or told of a potential link between self-affirmation and evaluations of threatening information (Study 3). Together, these studies suggest not only that affirmation processes can proceed without awareness but also that increased awareness of the affirmation may diminish its impact.

Keywords: self-affirmation, awareness, bias, self-esteem, defensiveness

In recent years, social psychology studies have shown that experimental self-affirmations can have intriguing consequences. "Self-affirmation" is the process by which people maintain a sense of self-integrity, that is, a perception of themselves as globally moral, adequate, and efficacious when they confront threats to a valued self-image (Steele, 1988). Self-affirmation manipulations typically involve writing about an important value or completing a value-relevant scale, with several notable effects. African American middle school students who wrote one or two essays about an important value at the beginning of the academic term ended up earning higher grades at the end of the term, attenuating the racial achievement gap with European American students in the same course (Cohen, Garcia, Apfel, & Master, 2006). College students who completed a scale concerning a personally important value had lower levels of the stress hormone cortisol when they underwent a stressful public speaking exercise (Creswell et al., 2005) and had reduced levels of epinephrine, a marker of sympathetic nervous system activation, on the day of their most stressful midterm examination (Sherman, Bunyan, Creswell, & Jaremka, in press). And women who were heavy alcohol consumers and wrote an essay about an important value were more open to health information linking alcohol use and negative health outcomes, with some effects persisting a month later (Harris & Napper, 2005). These studies are among a few of the striking examples from a literature showing that experimentally induced self-affirmations can reduce threat, stress, and defensiveness (see also

David K. Sherman and Debra P. Bunyan, Department of Psychology, University of California, Santa Barbara; Geoffrey L. Cohen and Julio Garcia, Department of Psychology, University of Colorado, Boulder; Leif D. Nelson, Haas School of Business, University of California, Berkeley; A. David Nussbaum, Department of Psychology, University of Waterloo, Ontario, Canada.

This research was supported by National Science Foundation Grant 0720429 to David K. Sherman. We would like to thank Kelly Hogan, Sarah LaForge, Allison Lunt, and Mary Prenovost for serving as research assistants; Matt Dunaj, Breeana Lancaster, and Andi Reed for serving as coders of the data; and Shelly Gable and Eric D. Knowles for statistical consultation. We would also like to thank Anthony Bastardi, Jeff Galak, Jane Risen, and Lee Ross for their contributions to this project and Thai Chu, Cynthia Gangi, Lisa Jaremka, Heejung Kim, Zoe Kinias, Wesley Moons, and Nate Way for commenting on earlier versions of this article.

Correspondence concerning this article should be addressed to David K. Sherman, Department of Psychology, University of California, Santa Barbara, Santa Barbara, CA 93106-9660. E-mail: david.sherman@psych.ucsb.edu

Journal of Personality and Social Psychology, 2009, Vol. 97, No. 5, 745–764
 0022-3514/09/$12.00 DOI: 10.1037/a0015451

Adams, Tormala, & O'Brien, 2006; Crocker, Niiya, & Mischkowski, 2008; Fein & Spencer, 1997; Schmeichel & Martens, 2005; for a review, see Sherman & Cohen, 2006).

Although self-affirmations in these studies have produced generally beneficial effects, several studies have demonstrated theory-predicted negative effects of self-affirmation under theoretically specified conditions, such as decreased open-mindedness in negotiation (Cohen et al., 2007, Study 3) and shallower information processing (Briñol, Petty, Gallardo, & DeMarree, 2007; see also Galinsky, Stone, & Cooper, 2000). Nevertheless, given some of these recent findings, there may be a strong sense that implementing self-affirmation widely, broadly, and frequently could yield, on average, positive effects. However, although affirmation interventions can have large-scale effectiveness, we propose that they should also adhere to small-scale subtlety. The self-affirmation process is not a panacea for stress, threat, and defensiveness, but rather, it plays a critical role in the operation of the "psychological immune system" that people use to protect the self when it is threatened (Gilbert, Pinel, Wilson, Blumberg, & Wheatley, 1998).

In keeping with that reasoning, we focus on a critical feature of the psychological immune system, namely, that people are unaware of it when it is operating. We suggest that heightened awareness of self-affirmation processes may actually be inimical to the positive outcomes documented above. In particular, we explore three facets of awareness in the affirmation process and advance two notions. First, people are generally unaware of the self-affirmation process in themselves; specifically, self-affirmation effects can proceed even when people are (a) unaware of having been self-affirmed, (b) unaware that a given self-affirmation manipulation affected their self-worth, and/or (c) unaware of a given self-affirmation's subsequent effect on their judgment and behavior. Second, to the extent that people are aware of these links in the psychological causal chain, affirmation effects may be attenuated.

Overview of Self-Affirmation Theory

Self-affirmation theory begins with the premise that people are motivated to maintain the perceived worth and integrity of the self and examines how people respond to information and events that threaten a valued self-image (Steele, 1988; see also Aronson, Cohen, & Nail, 1999; McQueen & Klein, 2006; Sherman & Cohen, 2006). Everyday life offers numerous potential psychological threats, whether it is negative feedback on an exam or health information implying that past behaviors put one at risk of disease. Gilbert, Wilson, and colleagues have theorized that the psychological immune system initiates protective adaptations under such impending threats, much like the response of the actual immune system under threat from pathogens (Gilbert et al., 1998). These psychological adaptations include the cognitive strategies and perceptual distortions that lead people to construe situations as less threatening to personal worth and well-being. These adaptations vary widely in strategy and in consequence.

For example, people respond to failure at times by attributing it externally (Miller & Ross, 1975) or by simply disidentifying with the threatened domain, sustaining self-worth but forestalling self-improvement (Major, Spencer, Schmader, Wolfe, & Crocker, 1998; Nussbaum & Steele, 2007; Steele, 1997). People view contradictory information through the prism of their ideology and may place greater scrutiny on information inconsistent with their prior beliefs (Lord, Ross, & Lepper, 1979). When health information suggests personal risk, people may challenge the information rather than plan to change their risky behavior (Ditto & Lopez, 1992; Kunda, 1987), contributing to a generally unrealistically optimistic view about personal health (Weinstein & Klein, 1995). All of these defensive judgments can help maintain the perceived integrity of the self by reducing potential threats.

Self-affirmation theory proposes that people seek to maintain a *global* sense of self-integrity rather than their perceived worth in specific domains or particular situations (Steele, 1988). Thus, if people can affirm an unrelated domain of self-worth, they will have less need to rationalize away threatening information (Sherman & Cohen, 2006; Steele, 1988). Indeed, many studies have demonstrated that affirming an alternative domain of self-worth can reduce defensive processing of threatening information (Cohen, Aronson, & Steele, 2000; Correll, Spencer, & Zanna, 2004; Fein & Spencer, 1997; Jacks & O'Brien, 2004; Reed & Aspinwall, 1998; Sherman, Nelson, & Steele, 2000). For example, people are more accepting of personally relevant, highly threatening health information, and intend to change their behavior accordingly, when they engage in a self-affirmation exercise, such as reminding themselves of their relationships with friends, their creativity, or another important personal value (Harris, Mayle, Mabbott, & Napper, 2007; Harris & Napper, 2005; Reed & Aspinwall, 1998; Sherman et al., 2000). When people reaffirm self-integrity through affirmations of this kind, counterattitudinal information is rendered less threatening to global self-integrity and individuals tend to respond without resorting to defensive biases (Sherman & Cohen, 2002).

Awareness and Self-Affirmation

Self-defensiveness and self-affirmation are both coping strategies associated with the same psychological immune system (see also Tesser, Crepaz, Collins, Cornell, & Beach, 2000). This system operates outside of awareness, or, as Gilbert et al. (1998) put it, "one of the hallmarks of the psychological immune system is that it seems to work best when no one is watching, and when its operations are explicitly scrutinized, it may cease functioning altogether" (p. 619; see also Wilson, 2002). Moreover, they argued that people are generally unaware of the influence of this psychological immune system, a phenomenon they term *immune neglect*, and which we believe is operating in the process of self-affirmation as well.

Yet, when considering self-affirmation, there are at least three aspects of awareness that could influence threat responses. First, people may or may not be aware that they are activating an important domain of self-worth. Prototypical self-affirmation manipulations—filling out a value-relevant scale or writing a value-relevant essay—necessarily require conscious awareness of the stimuli. Yet if self-affirmation effects could be observed with minimal conscious awareness of the affirming stimuli, then this would suggest that activities and actions undertaken to restore self-integrity could occur outside of deliberative awareness. We base this logic partly on Bargh's (1990) auto-motive hypothesis, which posits that goals can be activated and acted on completely outside of awareness. If people could exhibit typical self-affirmation effects in response to threats to self-integrity when they are not aware of the affirming stimuli (by, for example, using value-relevant sentence-unscrambling tasks), then this would suggest that the self-affirmation process may be part of this auto-motive system.

Second, people may or may not be aware that the self-affirmation manipulations buttress self-worth. Self-affirmation research has not examined whether people are aware of the effects of the standard affirmation manipulations. There is little reason to suspect that participants are aware that the affirmation manipulation is designed to buttress self-worth, as typically these tasks are presented in a manner that disguises their purpose to the participant. In this manner, the way in which social psychological research has tested the effects of self-affirmation stands in contrast to that of the self-help literature on "self-validation," in which the objective of the exercise is clear: "Optimal affirmations are first-person, present tense, optimal statements you implant in your mind to affirm the best results" (Glickman, 2002, p. 97). In the self-help literature, there is thus the expectation that affirmations and positive thinking should prove beneficial even, and perhaps especially, when individuals are aware of the rationale for engaging in them. By contrast, we suggest that self-affirmation may be most effective when people are unaware of its esteem-buffering effects.

One reason that more explicit affirmations that attempt to boost an individual's self-esteem may be less effective than more subtle affirmations is that the more explicit attempts could cause people to focus on the extrinsic benefits of the affirmational activity by leading them to think that the activity is done as a means to feel better about the self. When people explicitly pursue happiness, or explicitly strive to boost their self-esteem, such acts can be self-defeating and lead to less happiness and reduced self-esteem (Crocker & Park, 2004; Schooler, Ariely, & Loewenstein, 2003). Consistent with this notion, research has found that self-affirmations that focus on intrinsic aspects of the self (such as how personal values makes one feel) are more successful at reducing defensive threat responses than self-affirmations that focus on explicit aspects of the self (such as how important values lead to specific benefits; Schimel, Arndt, Banko, & Cook, 2004).

Third, people may or may not be aware that self-affirmations influence their subsequent judgments or performance (regardless of their awareness of their impact on their self-worth). In a typical self-affirmation study, participants believe they are taking part in two separate experiments, one involving personal values (the affirmation manipulation) and the other involving the evaluation of information (e.g., a threatening task or message; Cohen et al., 2000; Sherman et al., 2000; see McQueen & Klein, 2006, for a review of affirmation manipulations). This methodology effectively eliminates participant awareness of the connection between the two tasks. What would happen if participants were merely made aware of the possible connection between the two tasks? It may be that simply being made aware of a potential link between the affirmation and subsequent judgment might affect the impact of that affirmation.

In particular, some stimuli influence behavior only when people are unaware of the potential influence of the stimuli. Schwarz (2004) has proposed a model of metacognition to explain the role of awareness in implicit theories. People have strong implicit theories about how their thoughts are influenced by their environment, and when those potential influences are made salient, people try to correct for them (Wegener & Petty, 1995). One of the strongest implicit theories people possess is an illusion of objectivity, that is, that they perceive information in a bottom-up, rational manner (Pronin, Gilovich, & Ross, 2004; Ross & Ward, 1996). When external influence is made salient, people adjust their assessments accordingly. For example, phone survey respondents report being happier on sunny days than on cloudy days. However, if respondents first described the weather (cloudy or sunny), people corrected for the mood influences of weather, and there were no differences in happiness (Schwarz & Clore, 1983). Findings like these suggest that incidental external events (e.g., weather) can be influential when they escape conscious awareness. However, if the same external event is highlighted in some way, this increased awareness eliminates the influence of the external event. Awareness negates impact because, it seems, that people have lay theories connecting weather to mood, and when these theories are made salient, they try to correct for them (Wegener & Petty, 1995).

We contend that an effective self-affirmation, like the weather in the study described above, requires some lack of awareness. That is, people may be aware that a self-affirmation increases perceptions of self-integrity (Steele, 1988)[1] and that feeling better about oneself could affect subsequent judgments, just as they are aware that the weather could impact their mood. Moreover, drawing attention to the esteem-boosting functions of the self-affirmation, like drawing attention to the weather, could attenuate the impact of the manipulation because people may correct for the perceived influence of these factors (Wegener & Petty, 1995). Self-affirmation manipulations, then, may operate outside of awareness in an analogous manner to the sentence-unscrambling primes used in many social psychology studies (e.g., Bargh, Chen, & Burrows, 1996; Chartrand & Bargh, 1996; Epley & Gilovich, 1999; Kay & Ross, 2003) or in an analogous manner to the many incidental primes that affect people without their awareness in everyday life.

Awareness of a possible influence such as a self-affirmation manipulation may moderate the effect of the influence because people may want to avoid being influenced by such an incidental feature of the environment (Martin, 1986). The notion that their judgments or performance were influenced by something as seemingly unconnected to the task as a values affirmation may be inconsistent with people's implicit theories of how their thoughts and behaviors are influenced by the environment (Schwarz, 2004) and, more generally, with their views of themselves as rational information processors (Ross & Ward, 1996). Hence, affirmations that heighten awareness of the process could produce opposite effects of the typical self-affirmations in the same way that blatant primes produce opposite effects when compared to more subtle primes (Stapel, Martin, & Schwarz, 1998).

In the present article, we test five interrelated hypotheses about the role of awareness in the self-affirmation process:

1. The self-affirmation process can occur even when the self-affirmation is delivered outside of deliberative awareness via implicit, self-affirming primes.

2. Individuals are generally unaware that self-affirmation manipulations affect their performance or judgments in potentially threatening situations.

3. As an extension of Hypothesis 2, perceived influence of self-affirmation should *not* be positively correlated with actual influence. This would be the case if people have

[1] To assess this greater perception of self-integrity, in Studies 1 and 2, we utilize a self-integrity scale designed to capture feelings of moral and adaptive adequacy.

little if any introspective access to the causal effects of self-affirmation on their judgment and behavior.

4. Experimentally manipulating awareness—by informing people that a self-affirmation exercise is designed to enhance self-esteem—will attenuate the standard self-affirmation effects.

5. Merely suggesting the possible impact of self-affirmation by informing people of a general influence of the affirmation exercise on the subsequent activity, without specifying the direction of that influence, will attenuate the standard self-affirmation effects.

In three studies, we directly examine these hypotheses. In Study 1 we examine whether the self-affirmation process can occur when the manipulation is delivered outside of deliberative awareness in the form of an implicit affirmation (Hypothesis 1). We also directly examine whether people report being aware of the impact of the implicit self-affirmations on their judgments (Hypothesis 2) and whether perceived influence is correlated with actual influence (Hypothesis 3). Study 2 examines whether informing participants that the manipulation is designed to "boost self-esteem" reduces the effectiveness of the affirmation (Hypothesis 4) and measures again the correspondence between perceived and actual influence of the affirmation procedure (Hypotheses 2 and 3). Finally, Study 3 considers whether merely linking the manipulation with the dependent measure—thereby introducing the potential influence of affirmation without specifying a direction of that influence—would suffice to reduce the manipulation's effectiveness (Hypothesis 5). Study 3 also provides additional evidence of the effectiveness of implicit self-affirmations (Hypothesis 1).

We present these three experiments as explorations in self-affirmation research that address the role of awareness in the integrity-maintenance process. We provide convergent evidence for the affirmation process in three domains: performance on a potentially threatening math test (Study 1), the evaluation of group-identity-threatening information (Study 2), and the assessment of personal health risks (Study 3).

Study 1

In previous self-affirmation research, participants have been aware of the self-affirming stimuli but not, we contend, of their potential influence on the outcome measure. They were aware that they were completing a values scale or writing an essay about an important value but unaware that this might affect their responses to the subsequent task. What are the effects of presenting self-affirmation at a more implicit level, that is, at a level at which the self-affirming thoughts are activated using a more implicit procedure?

Past research has shown that nonconsciously activated goals can have the same effects as consciously activated goals (Chartrand & Bargh, 1996). Simply unscrambling goal-related sentences can replicate effects otherwise thought to operate only through more explicit goal pursuit. Analogously, Study 1 employs a sentence-unscrambling priming procedure to affirm individuals and compares this implicit affirmation to a neutral prime condition. If the implicit affirmation demonstrates similar effects as a standard self-affirmation, then it would suggest that self-integrity goals can be satisfied nonconsciously.

Study 1 used a manipulation designed to operate below participant deliberative awareness, but we also measured how people thought the manipulation might influence their subsequent behavior. These measures allowed us to generally assess self-reported awareness as well as look at the relationship between such awareness and effectiveness of the affirmation. We emphasized to participants that their answers were confidential and would not be linked to their names to minimize social desirability pressure in order to assess, as best we could, the extent to which they believed that the affirmation—as well as other potential factors—was influencing them. In Study 1 as well, we introduced a new measure to check the manipulation of self-affirmation (cf. McQueen & Klein, 2006), a Self-Integrity Scale (Cohen, Garcia, & Sherman, 2009).

Study 1 examines performance on standardized tests after failure. In recent research by Cohen et al. (2006; Cohen, Garcia, Purdie-Vaughns, Apfel, & Brzustoski, in press), a self-affirmation intervention was designed to alleviate the stress arising from threats to students' social identity. For ethnic minority students in general and for female students in math and science contexts, such threats arise from the concern over confirming a negative stereotype about the intellectual ability of one's racial or gender group (Steele, 1997; Steele & Aronson, 1995; Steele, Spencer, & Aronson, 2002). By increasing stress, this concern can undermine performance (O'Brien & Crandall, 2003). By contrast, self-affirmations have been shown to reduce stress (Creswell et al., 2005; Sherman et al., in press) and increase performance among individuals in threatening academic environments (Cohen et al., 2006; Martens, Johns, Greenberg, & Schimel, 2006).

An examination of the results of Cohen et al. (2006) found that the affirmation appeared to buffer minority students against the impact of early poor performance. In the no-affirmation condition, minority students displayed a downward trend in performance; the more their performance dipped early in the term, the worse their performance was later. The self-affirmation seemed to break this recursive process whereby failure led to worsening performance over time. That is, affirmation-treated minority students did not display as steep a downward trend in performance over time, and there was no correlation between any early dip in their performance and their performance later in the term. This occurred because the affirmation buffered people psychologically against the impact of poor performance, making them more resilient and likely to perform better on subsequent tasks (Cohen & Garcia, 2008).

We sought to examine this recursive process in the laboratory. Thus, in the current study we aimed to first instantiate an academic self-threat (by having participants take an extremely difficult math test) and then to measure performance on a different, less difficult quantitative exam. Furthermore, conceptually replicating procedures used in past research (e.g., Davies, Spencer, Quinn, & Gerhardstein, 2002), the latter measure offered people a chance to either work on a threat-relevant task (i.e., a math exam) or a threat-irrelevant task (i.e., a verbal test). If participants are under self-threat, it has been shown that they are more likely to avoid the domain of threat, in this case, math (Davies et al., 2002).

Prior to the extremely difficult math test, however, half of the participants had completed an implicit affirmation, where they un-

scrambled sentences relevant to an important value, while the remaining participants unscrambled neutral sentences. It was predicted that the implicit self-affirmation would enable people to cope better with the failure of the first test, which would serve as the threat, and help them on the second, moderately difficult math test.

A secondary prediction concerns the moderating role of domain identification; we predicted that students highly identified with math would be most threatened in the absence of the affirmation and most buffered by the affirmation. By definition, students more ego-invested in math should be more threatened by poor performance in math and, consequently, have the greatest potential to benefit from the affirmation (see also Aronson, Lustina, et al., 1999; Sherman, Kinias, Major, Kim, & Prenovost, 2007).

Method

Participants and Design

During a pretest, participants ranked five values in terms of their personal importance: social issues, art, science, politics, and religion. These values were taken from the Allport, Vernon, and Lindzey (1960) study of values and were used to assign participants in the implicit affirmation condition. The 48 participants had a mean age of 20.47 years ($SD = 2.42$), consisted of 19 men and 29 women, and included 28 European Americans, 2 African Americans, 10 Latinos, 4 Asian Americans, and 4 other/missing data.

Participants were randomly assigned to one of two conditions: affirmation prime or neutral control prime. One participant did not answer any questions on the first test and so was eliminated from the sample. During pretest (prior to the experimental session), participants also indicated on 9-point scales how skilled they were at math and how important doing well on math standardized tests was to them. These two items ($\alpha = .75$) formed an index of math identification.

Procedure

Affirmation manipulation. Participants were told that the researchers were developing various exercises and that they would be completing some language and some math exercises. They first completed the 10-min "language exercise" (which served as the priming procedure), a procedure that was modeled after Bargh et al. (1996) and designed to implicitly prime either self-affirming thoughts or neutral, nonaffirming thoughts. Participants unscrambled 30 sentences by removing one word from a list and making a sentence with the remainder. Those in the affirmation prime condition were given a sentence-unscrambling task that corresponded to the value that they had indicated in pretest as being most important. For example, affirmation participants who indicated art as their most important value unscrambled 15 sentences containing art-relevant primes, such as "colors blended the pretty well" into "the colors blended well." Participants in the neutral prime condition unscrambled 30 neutral primes, unrelated to any particular value (e.g., "couches were both flower red" could be unscrambled as "both couches were red.")

Math and verbal tests. After completing the prime, participants took a math test that was described as diagnostic of their math ability. The test consisted of 30 problems drawn from the most difficult items on the Graduate Record Examination (GRE) math test (as in Spencer, Steele, & Quinn, 1999). This test was chosen because it would be extremely difficult and presumably threatening to all of the college student participants. Hence we will refer to it as the *threat-inducing math test*. Participants had 15 min to work on the test.

After the 15 min had elapsed, participants completed a number of questionnaires about the test and their interests in math. Participants then received 15 math problems from the SAT (the *moderately difficult math test*) and 15 verbal problems from the SAT (the *moderately difficult verbal test*). They were told that they could work on either set of problems that they wanted to for the next 15 min. The math and verbal tests were presented on alternating pages (see also Davies et al., 2002).

Math pretest. We first wanted to confirm that the threat-inducing test was perceived as more difficult than the subsequent test we used as our critical measure. Undergraduates ($N = 14$, 7 women, 7 men) rated both math tests on 9-point scales anchored at 1 (*not at all difficult*), 5 (*moderately difficult*), and 9 (*extremely difficult*) and judged the threat-inducing math test to be much more difficult than the domain-relevant math test ($M = 6.56$, $SE = 0.31$ vs. $M = 4.62$, $SE = 0.51$), paired $t(13) = 3.41$, $p = .005$. Furthermore, participants rated the same test to be a more stressful test to take "over 15 minutes as part of a psychology experiment" ($M = 7.36$, $SE = 0.48$ vs. $M = 4.85$, $SE = 0.65$), paired $t(13) = 3.83$, $p = .002$. In sum, the two tests differed in difficulty, which we predicted would manifest itself in threat that would be interpreted in light of the affirmation manipulation (cf. Cohen et al., 2006).

Manipulation check. Participants then completed a number of posttest questionnaires, including a manipulation check designed to assess their perceptions of their self-integrity (Cohen et al., 2009). The scale contained 8 items ($\alpha = .84$); participants indicated their agreement with each item on a scale from 1 (*strongly disagree*) to 7 (*strongly agree*). The items include statements designed to assess a feeling of general moral and adaptive adequacy (Steele, 1988), such as "I feel that I'm basically a moral person" (see Appendix for the scale).

Awareness probe. Finally, participants completed a detailed probe of awareness. They first reported in open-ended fashion those factors they thought could have contributed to their performance on the math test. We then listed 12 potential factors and asked participants to rate the extent to which each factor influenced their math performance. Those factors were as follows: your math ability, your effort on the test, your personal background, your beliefs, your attitudes, your familiarity with math problems, the first language task that you completed, your reasoning ability, your self-esteem, your mood, the gender of the experimenter, and the time of day. Participants rated each of those factors on a scale from 1 (*did not contribute at all*) to 9 (*contributed a great deal*) in terms of how much each factor contributed to their math performance.

Results

Preliminary Analyses and Self-Integrity Manipulation Check

As a check on random assignment we examined whether the identification premeasure varied across conditions. We conducted a multivariate analysis of variance (MANOVA) on the two items that composed the measure. The MANOVA revealed no difference as a function of condition, $F(2, 44) = 0.46$, $p = .63$. Furthermore, we conducted a one-way analysis of variance (ANOVA) on the two-item composite, and the identification premeasure did not vary

across conditions, $F(1, 45) = 0.01$, $p = .93$, indicating success of randomization. Overall, participants were highly identified with math and saw standardized math tests as important (composite $M = 6.18$, $SE = 0.23$, on the 9-point scale).

In terms of the manipulation check, participants in the implicit affirmation condition scored higher on the Self-Integrity Scale ($M = 5.75$, $SE = 0.14$) than did those in the control condition ($M = 5.21$, $SE = 0.15$), $F(1, 45) = 6.61$, $p = .014$, suggesting that the implicit affirming primes increased participants' perception of their own self-integrity. Next, we examine whether the implicit affirmation prime, relative to the neutral prime, affected individuals' performance.

Effect of Self-Affirmation on Performance

We conducted a one-way MANOVA to examine the effect of prime (implicit affirmation vs. neutral) on the performance for the three different tests (threat-inducing math test, moderately difficult math test, moderately difficult verbal test). We predicted that, following the threat of a very difficult math test, the implicit self-affirmation would improve performance on a moderately difficult math test. Table 1 indicates the mean number of problems attempted, mean number correct, and accuracy (the number correct divided by the number attempted) for the three tests.[2]

The threat-inducing math test was, in fact, quite difficult. Participants overall attempted to answer fewer than half the problems ($M = 11.62$, $SE = 1.45$) and answered very few correctly ($M = 3.17$, $SE = 0.34$). Furthermore, these results were not influenced by the affirmation manipulation, $F(1, 45) = 0.59$, $p = .45$, for number attempted; $F(1, 45) = 1.63$, $p = .21$, for number correct. Thus, the implicit self-affirmation manipulation had no effect on the initial, extremely difficult math test. It also had no effect on the domain-irrelevant, moderately difficult verbal task.

However, consistent with our predictions, the implicit self-affirmation did affect how people performed on the subsequent, moderately difficult math test. Table 1 (bottom) displays the means for number attempted, number correct, and accuracy by condition for the math and verbal problems. Significant main effects emerged on the number of math problems correct and math problem accuracy. Those in the implicit affirmation condition answered more math questions correctly ($M = 1.88$, $SE = 0.27$) than did those in the control condition ($M = 0.73$, $SE = 0.29$), $F(1, 45) = 8.64$, $p = .005$. Furthermore, of the math problems attempted, a greater percentage were answered correctly by those in the implicit affirmation condition ($M = 0.39$, $SE = 0.06$) than among those in the control condition ($M = 0.17$, $SE = 0.06$), $F(1, 44) = 8.00$, $p = .007$. There was no main effect of condition on the number of moderately difficult math problems attempted, as those in the implicit affirmation condition ($M = 4.36$, $SE = 0.59$) did not differ from those in the control condition ($M = 4.09$, $SE = 0.63$), $F(1, 45) = 0.10$, $p = .77$. Participants who completed the implicit affirmation performed better on the moderately difficult math task after failing at the threat-inducing math task.

Next, we examined whether identification with math would moderate the effect of the affirmation. To simplify analysis, we computed change scores in math accuracy by subtracting accuracy on the first math test from accuracy on the second math test (with accuracy for both tests being calculated by the number correct divided by number attempted). Thus positive numbers represent an improvement in performance whereas negative numbers indicate a decrement in performance. We regressed this change score on the mean-centered identification measure, affirmation condition (contrast coded control condition = –1, implicit affirmation condition = 1), and the interaction. There was a significant effect of condition, $\beta(42) = .398$, $p = .005$, indicating that the affirmed participants improved whereas the nonaffirmed did not. However, this was qualified by an Identification × Condition interaction, $\beta(42) = .280$, $p = .044$. We interpreted the interaction by plotting the predicted means at 1 *SD* above and below the mean on the identification measure, following Aiken and West (1991). As can be seen in Figure 1, for those who were not identified with math, the affirmation manipulation did not have much effect, as both those in the no-affirmation condition (estimated $M = -0.04$) and the implicit affirmation condition (estimated $M = 0.02$) did not differ much from pretest to posttest, $\beta(42) = .101$, $p = .61$. However, for those highly identified with math, in the absence of affirmation, there was a decrement in performance (estimated $M = -0.29$), whereas those who implicitly affirmed themselves had an improvement in performance (estimated $M = 0.16$), $\beta(42) = .690$, $p = .001$. In sum, the implicit self-affirming primes enabled students to perform better after threat.

Finally, we examined the relative proportion of verbal and math problems attempted during the key testing phase. We computed the proportion of math items attempted by dividing the number of math items attempted by the total number of items attempted (math + verbal). Thus, a score of 1.00 indicates that a person

Table 1
Performance on Math and Verbal Tests in Study 1 as a Function of Affirmation Condition

Test	Control	Implicit affirmation
Threat-inducing math		
Attempted	12.46 (1.50)	10.88 (1.41)
Correct	3.50 (0.35)	2.88 (0.33)
Accuracy	0.33 (0.04)	0.30 (0.04)
Moderately difficult verbal		
Attempted	8.09 (0.71)	8.16 (0.64)
Correct	2.55 (0.41)	2.96 (0.39)
Accuracy	0.33 (0.09)	0.36 (0.04)
Moderately difficult math		
Attempted	4.09 (0.63)	4.36 (0.59)
Correct	0.73 (0.29)	1.88 (0.27)*
Accuracy	0.17 (0.06)	0.39 (0.06)*

Note. Accuracy is computed by number correct divided by number attempted for that type of problem. Data are means with standard errors reported parenthetically.
* $p < .05$ for condition difference between implicit affirmation and control conditions.

[2] Three participants did not attempt any math problems on the second test, and so we assigned them 0% accuracy in the primary results to preserve degrees of freedom. All 3 participants were in the implicit affirmation condition, so, this is a conservative decision, as these participants' data depress the scores in the implicit affirmation condition at Time 2. Also, 1 person (in the control condition) attempted only one problem and received one correct, yielding 100% accuracy. Because this was an outlier (greater than 3 *SD*s from the group mean), this score was eliminated from percentage analyses (but not from analyses of absolute number of problems correct and attempted), resulting in slightly different degrees of freedom.

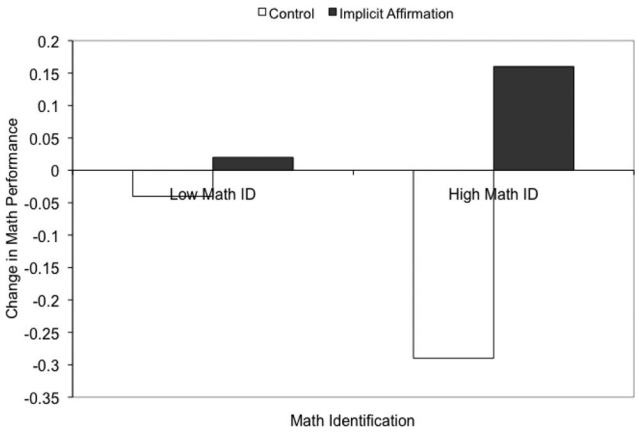

Figure 1. Change in performance (percent correct) as a function of math identification (ID) and affirmation condition in Study 1. Values are plotted at 1 *SD* above and below the mean of math identification. Positive numbers indicate improvement from first test to second test, whereas negative numbers indicate decreased performance.

attempted all math problems and no verbal problems and a score of 0.00 indicates that a person attempted all verbal problems and no math problems. We regressed this measure on the mean-centered identification measure, affirmation condition (contrast coded control condition = –1, implicit affirmation condition = 1), and the interaction. As the measure was not normally distributed, we took the square root of the percentage. As participants varied widely in how many problems they attempted, the analysis also controlled for number of problems attempted from the first more difficult math test.[3] There was no main effect of condition, $\beta(42) = -.05$, $p = .73$, as those in the affirmation condition attempted the same proportion of math problems as those in the no-affirmation condition. There was also a main effect for math identification, as the more math-identified students attempted a greater proportion of math problems, $\beta(42) = .297$, $p = .05$. However, this was qualified by an Identification × Condition interaction that approached significance, $\beta(42) = .273$, $p = .081$.

When participants were affirmed, their math identification positively predicted the proportion of math problems attempted, $\beta(42) = .571$, $p = .021$. We estimated means at 1 *SD* above and below the mean on the identification measure. The more identified students chose to work on a greater proportion of math problems (estimated $M = 0.43$) than did the less identified math students (estimated $M = 0.21$) when they were affirmed. By contrast, in the absence of an affirmation, identification did not predict choice of problem, $\beta(42) = .024$, $p = .89$, as there was no difference between weakly (estimated $M = 0.33$) and strongly (estimated $M = 0.34$) math-identified students. The affirmation enabled math-identified students to attempt as well as to perform better on more math problems of moderate difficulty.

Awareness of the Affirmation

Participants were generally unaware of the implicit affirmation. Two coders, unaware of condition assignment, examined the open-ended responses to the probe that asked what factors influenced their performance. No participant cited the language exercise (i.e., the sentence-unscrambling task) as directly influencing his or her performance. Only 1 participant mentioned the sentence-unscrambling task at all, noting that the sentence-unscrambling task was much easier in comparison to the math test (e.g., "the first language test was completely easy, low level"). Although most of the factors that participants listed concerned math (46% concerned their own math ability, and 14% concerned the difficulty of the math test), 26% referred to some internal quality of the participant (the remaining 14% referred to external other factors such as noise). Two coders classified the valence of the "internal" category into one of three categories: positive, negative, or neutral/valence unspecified. If any participant mentioned factors related to self-affirmation, such as "I felt good about myself" it would be coded as positive. Overall, there was 83% agreement amongst the coders, and they resolved the discrepancies via discussion (all discrepancies were along the neutral–negative distinction). Of the internal other comments, 83% were coded as negative (e.g., "I am not taking math classes so my mind isn't as sharp as usual"), 17% were coded as neutral (e.g., "state of mind"), and 0% were coded as positive. In sum, the open-ended responses yielded no evidence to suggest that participants were aware of the implicit affirmation's impact on their performance.

Next, we examined how participants rated the 12 potential performance-influencing factors. Because we were interested in the relative subjective importance of the different potential factors of influence, we standardized the 12 factors within person and conducted a MANOVA with the 12 factors as the dependent measures and with condition as the independent variable (1 person left 1 of the 12 factors blank, and we assigned the overall group mean to fill that missing cell). The multivariate analysis revealed no main effects of condition, $F(12, 34) = 0.67$, $p = .77$ (see Figure 2). The language exercise (the self-affirmation manipulation) was judged as one of the least important factors (on average rated #9 out of 12; $M = -0.39$), and there was no main effect of condition ($M = -0.42$, $SE = 0.16$, for implicit affirmation condition; $M = -0.35$, $SE = 0.16$, for neutral condition), $F(1, 45) = 0.13$, $p = .73$.[4] In sum, there was no evidence for participants' awareness of the impact of the affirmation at the mean level. Next, we examine evidence for awareness of the impact of the affirmation at the

[3] The number of problems attempted was highly correlated between the first math test and the second test (combined math + verbal problems attempted), $r(47) = .48$, $p = .001$. However, the number attempted on the difficult math test did not differ across conditions, $F(1, 45) = 0.59$, $p = .45$.

[4] Examination of the individual items revealed only one main effect that approached significance on the item "mood," as those in the implicit condition ($M = 0.46$, $SE = 0.16$) thought mood influenced them more than did those in the control condition ($M = 0.07$, $SE = 0.17$), $F(1, 45) = 3.00$, $p = .09$. We also examined whether there was any effect of sex or any Sex × Condition interactions on all analyses. In terms of the awareness probe, there were two significant main effects. Women ($M = -0.60$, $SE = 0.13$) thought they were less influenced by the language activity (the sentence unscrambling) than men ($M = -0.19$, $SE = 0.16$), $F(1, 42) = 4.11$, $p = .049$. There was no interaction between sex and condition. Women also thought they were more influenced by their familiarity with the problems ($M = 0.97$, $SE = 0.15$) than men ($M = 0.49$, $SE = 0.17$), $F(1, 42) = 4.37$, $p = .043$. In terms of performance effects, there were no main effects of sex and no interactions between sex and condition on the different measures of performance. There was also an insufficient number of participants to examine ethnicity as a factor.

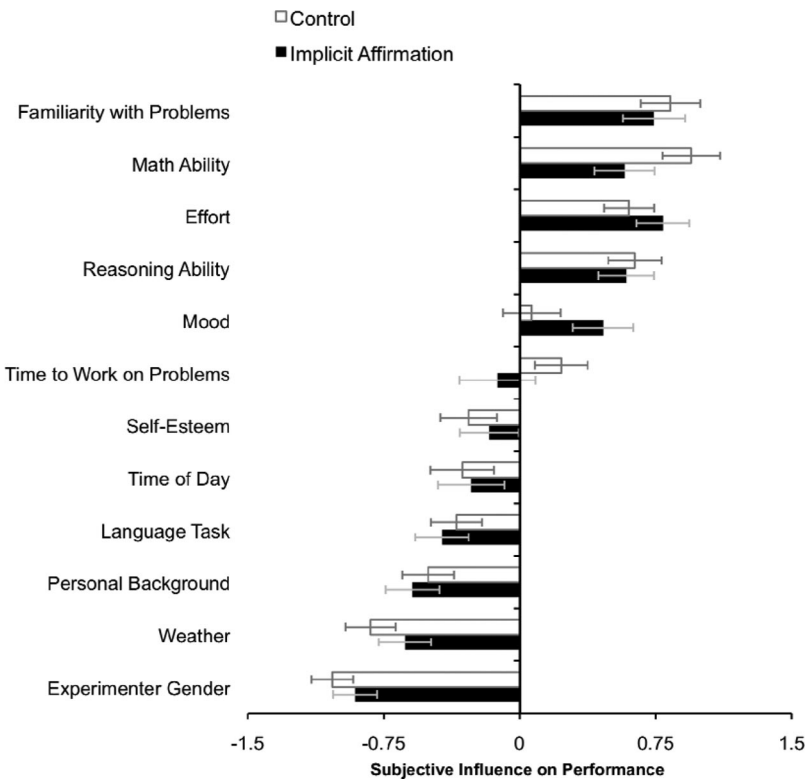

Figure 2. The subjective influence of 12 factors measured in Study 1. Higher numbers indicated that participants thought that factor was more influential. Error bars denote standard error. Note that "language task" referred to the affirmation manipulation.

correlational level by assessing whether participants' estimated influence correlated with the actual influence of the manipulation.

Relationship Between Awareness of Affirmation and Performance

We next examined whether participants' awareness of the manipulation influenced actual performance (change in performance from the first test to the second test). We regressed this change score on the mean-centered awareness measure, affirmation condition (contrast coded control condition = –1, implicit affirmation condition = 1), and the interaction. There was a significant effect of condition, $\beta(42) = .417$, $p = .003$, indicating again that the affirmed participants improved whereas the nonaffirmed did not. However, this was qualified by an Awareness × Condition interaction, $\beta(42) = -.288$, $p = .05$. We interpreted the interaction by plotting the predicted means at 1 *SD* above and below the mean on the awareness measure. For those in the no-affirmation condition, there was a nonsignificant difference between those low (estimated $M = -0.24$) and high (estimated $M = -0.08$) in awareness (i.e., the estimated impact of the language exercise), $\beta(42) = .27$, $p = .28$. However, for those in the implicit affirmation condition, increased perceived awareness of the impact of the implicit affirmation was associated with decreased benefit of the affirmation, as those who were low on awareness (estimated $M = 0.18$) did better than those who were high on awareness (estimated $M = -0.01$), $\beta(42) = -.31$, $p = .058$. In sum, the implicit self-affirming primes enabled students to perform better after threat, but awareness of the affirmation interfered with this effect.

Finally, we examined whether awareness and identification interacted and whether this varied by condition. We regressed this change score on the mean-centered awareness measure, the mean-centered math identification measure, affirmation condition (contrast coded control condition = –1, implicit affirmation condition = 1), and all possible two-way and three-way interactions. There was no significant three-way interaction, $\beta(38) = .17$, $p = .21$, and the only effects to emerge were those reported above: the significant main effect of condition, $\beta(38) = .44$, $p = .002$, the significant interaction between math identification and condition, $\beta(38) = .29$, $p = .035$, and an interaction between awareness and condition that approached significance, $\beta(38) = -.26$, $p = .081$. In sum, the relationships between awareness and condition and between identification and condition appear to be independently predicting performance.

Discussion

Study 1 found support for the three hypotheses it tested. First, supporting Hypothesis 1, Study 1 demonstrated the effectiveness of an implicit self-affirmation. Following a threat-inducing math test, a self-affirming sentence-unscrambling task improved performance on a subsequent, moderately difficult math test. Consistent with past research, the beneficial effects of the affirmation were stronger among those highly identified with the domain, the indi-

viduals who were otherwise most threatened (Aronson, Lustina, et al., 1999; Sherman et al., 2007). Participants who were highly identified with math performed best when self-affirmed and also chose to answer more problems in the threatened domain. Furthermore, consistent with previous findings (e.g., Cohen et al., 2006), the affirmation had a delayed performance effect. Performance was unaffected on the first, threat-inducing math task but was substantially improved on the subsequent domain-relevant math task. That is, the implicit self-affirmation appeared to stop a recursive process whereby failure may increase doubt that may in turn worsen performance.

In support of Hypothesis 2, participants thought that manipulation did not influence their performance, as indicated by their open-ended responses and their relatively low mean level perception of influence. However, there were individual differences in how much participants thought that they were influenced by the affirmation, and the regression analyses suggest that it was among those participants who were lowest in awareness that the implicit affirmation exerted the strongest beneficial effects on performance. That is, in the implicit affirmation condition, the more participants thought that the implicit affirmation could have influenced them, the worse their performance. This suggests, in support of Hypothesis 3, that awareness of the potential impact of the affirmation may attenuate its effectiveness.

Study 1 also raises new questions. Although an implicit affirmation was shown to influence behavior in much the same way as the standard affirmations have done (Cohen et al., 2006; see also Martens et al., 2006; Schimel et al., 2004), the study does not address whether people are aware of the influence of a standard affirmation, such as an essay writing manipulation, that requires greater deliberative awareness of the affirming stimuli. Furthermore, if people were made more aware of the affirmation process would that impair or reduce its effectiveness? The next two studies explore two possible aspects of the affirmation process that, when participants are experimentally made aware of, could attenuate the effectiveness of the affirmation: awareness that the affirmation could boost self-esteem (Study 2) and awareness of the affirmation's potential relationship to subsequent judgments (Study 3).

Study 2

If people knew that an affirmation could boost their self-esteem, would the affirmation lose its effectiveness? In Study 2, we compare the effects of a standard self-affirmation and a standard no-affirmation control condition to a condition in which participants were made aware of the potential self-affirmation influence. Participants in this awareness + affirmation condition were told that the writing activity was designed to increase their self-esteem. We designed this manipulation to reflect the possibility, supported by the self-integrity manipulation check findings in Study 1, that self-esteem was a plausible effect of the self-affirmation manipulation that participants could, in theory, be aware of. We predicted that once we raised awareness of the self-affirmation influence on self-esteem, the beneficial impact of the self-affirmation would be reduced because participants would see the activity as a means to an end, and thus, the affirmation would seem less important, undermining its self-affirming qualities. In addition, in Study 2, we examined whether people were aware of the influence of the standard self-affirmation manipulation by including the detailed awareness probes as in Study 1.

We examined these issues by asking fans of a professional sports team to evaluate information threatening to their sports-fan identity. Fans of the San Francisco Giants baseball team read information about Barry Bonds—perhaps the current player most emblematic of the Giants franchise at the time of the study—and his alleged steroid use. People evaluate identity-relevant information in a manner congenial to that identity. For example, capital punishment proponents and opponents view ambiguous evidence as supporting their particular side of the debate (Lord et al., 1979). In sports, fans and athletes interpret victories and defeats in a group-serving manner, attributing victory more than defeat to internal factors (Lau & Russell, 1980; see also Hastorf & Cantril, 1954). Moreover, self-affirmation attenuates these group-serving biases, suggesting that they stem from the need to protect an individual's identity (Cohen et al., 2007; Sherman & Kim, 2005).

The present study presented San Francisco Giants fans with potentially identity-threatening information about Barry Bonds. Bonds, who became the all-time Major League Baseball (MLB) home run leader in 2007 (shortly after this study was conducted), had been accused of using performance-enhancing drugs in many newspaper accounts, and most thoroughly, in the book *Game of Shadows* (Fainaru-Wada & Williams, 2006; see also Mitchell, 2007). Participants read an editorial suggesting that MLB should not celebrate Barry Bonds's accomplishments, a position likely at odds with any identity they held as a "Giants fan." In line with past research (Lord et al., 1979), we predicted that Giants-identifying individuals would be less open to the editorial. Specifically, we predicted that the more identified participants were with being a fan of the San Francisco Giants, the more critical they would be of the article. Moreover, we predicted that a standard self-affirmation manipulation would reduce this effect (Cohen et al., 2000) by eliminating the relationship between identification and openness to the report (as in Cohen et al., 2007). Beyond examining the impact of self-affirmation in a novel community sample in Study 2, we examine whether participants are aware of the impact of the standard affirmation (Hypothesis 2), whether perceived impact of the affirmation correlates with actual impact (Hypothesis 3), and whether experimentally increasing awareness that the affirmation could boost self-esteem would reduce the affirmation's effectiveness (Hypothesis 4).

Method

Participants and Design

Participants were recruited on public transportation (a commuter train) en route to a San Francisco Giants baseball game. Some participants were fans on the way to attend a game and some were not. Data collection took place in two waves, first in summer 2006 ($n = 34$) and then again in summer 2007 ($n = 57$) prior to Barry Bonds breaking the home run record; collection wave did not influence the results so we combined the samples. The participants had a mean age of 33.1 years ($SD = 16.2$), consisted of 52 men and 36 women, and included 65 European Americans, 5 African Americans, 6 Latinos, and 15 other/missing data (3 people declined to

report all demographics). Participants were randomly assigned to one of three conditions: self-affirmation condition, no-affirmation condition, and awareness + affirmation condition.

Procedure

Participants were offered $5 or a $5 gift card to complete a study on "spectator sports" during northbound train rides to San Francisco. Volunteers raised their hand and were given the materials. They were compensated when the materials were returned. The ride was advertised as serving the Giants games and was expected to have a large proportion of San Francisco Giants fans aboard. Participants learned that the study was about fans' views of themselves and evaluation of an essay about Barry Bonds and the home run record. The materials contained (in order) the Giants identification measures, the self-affirmation manipulation, a short essay about Barry Bonds, the dependent measures, and the awareness probe. The study lasted approximately 15 min.

Giants identification. Prior to the affirmation manipulation, participants indicated their favorite baseball team and the number of San Francisco Giants games, on average, that they attended per year. Participants also rated the personal importance of being a San Francisco Giants fan on a scale anchored at 1 (*not at all important*) and 9 (*extremely important*) and the extent to which they agreed with the statement, "Being a San Francisco Giants fan is an important reflection of who I am" on a scale anchored at 1 (*strongly disagree*) and 9 (*strongly agree*). These three measures (number of Giants games attended and the two self-report identification measures) were standardized and averaged ($\alpha = .75$) to form an index of Giants identification.

Self-affirmation manipulation. Participants next completed a version of a standard self-affirmation manipulation (McQueen & Klein, 2006; Sherman & Cohen, 2006; Sherman et al., 2007). First, they ranked 10 personal values (e.g., sense of humor, social skills, business/money, relations with friends) in order of importance and then completed a related writing exercise. Participants in the self-affirmation condition wrote three reasons that their most important value was important to them and then provided an example demonstrating its importance. Those in the no-affirmation condition wrote three reasons why their 9th-ranked value might be important to someone else and provided a relevant example. Finally, participants in the awareness + affirmation condition were given the same instructions as those in the self-affirmation condition, but for these participants the following statement appeared in large print directly above the space in which they wrote their essay: "The below writing activity is designed to make you feel better about yourself and to increase your self-esteem." This additional instruction was theorized to make participants aware that the self-affirmation could boost their self-esteem.

Persuasive essay. Participants all read a persuasive, one-page essay, purportedly written by a Bay Area sportswriter titled, "Let's Not Celebrate a Cheater" which argued that, because of his alleged steroid use, MLB should not commemorate Bonds when and if he passed the home run record. The article was drawn from various authentic articles and was in the format of an op-ed column featured in a sports section. The article included baseball statistics and evidence drawn from *Game of Shadows* consistent with the argument that Barry Bonds used performance-enhancing drugs and therefore should not be honored for his home run record. The crux of the article can be summed up in the below quotation:

> As Malcolm Gladwell (author of *Tipping Point*) put it, "*Game of Shadows* is a death sentence for Bonds. More to the point, it's impossible to read the book and accept that Bonds has a right either to the single season home-run record or, assuming he keeps playing, the career home run mark" (Gladwell, 2006).

The opinion piece concluded by arguing that due to his "obvious use of steroids, Bonds should not be celebrated or commemorated if and when he passes Hank Aaron."

Openness to the report. Participants rated statements about the essay and its author. Four key items formed the dependent measure assessing the participants' openness to the report. Participants rated the article's validity on a scale anchored at 1 (*not at all valid*) and 9 (*extremely valid*) and the author's objectivity and intelligence on two 9-point scales anchored at 1 (*not at all objective/intelligent*) and 9 (*extremely objective/intelligent*). Participants also indicated their interest in reading *Game of Shadows* on a similar 9-point scale. These four items ($\alpha = .74$) were averaged and combined into a single measure evaluating favorability toward the article. As a manipulation check, participants then completed the Self-Integrity Scale (see Appendix).

Awareness probe. The first wave of participants completed an open-ended probe for awareness presented as an empty sheet of paper with the prompt, "What factors do you think may have contributed to your thoughts and feelings in regards to the essay about Barry Bonds?" Both waves of participants also rated the extent to which 11 factors influenced their evaluation of the essay, including, for example, your knowledge of baseball, your identification with the Giants, and the first writing exercise you completed (about values). As in Study 1, we emphasized to participants that their answers were confidential and would not be linked to their names to minimize social desirability pressure in order to assess the extent to which they thought that various factors—including the self-affirmation manipulation—influenced them.

Results

Preliminary Analyses and Self-Integrity Manipulation Check

As a check on random assignment we examined whether the identification premeasure varied across conditions. We conducted a MANOVA on the three items that composed the measure. The MANOVA revealed no difference as a function of condition, $F(6, 166) = 1.13$, $p = 1.35$. Furthermore, we conducted a one-way ANOVA on the three-item composite, and the identification premeasure did not vary across conditions, $F(2, 88) = 0.58$, $p = .56$, indicating success of randomization. The sample of participants, on average, reported attending 4.85 baseball games per year ($SE = 0.94$; median = 2 games, mode = 2 games; the mean was inflated by 1 fan who reported attending 65 home games a year). Overall, 67% indicated that the Giants were their favorite team and the sample was lower than the midpoint (5) on both identification items. For the item assessing personal importance of being a San Francisco Giants fan, the overall mean was 4.22 ($SE = 0.30$), and for the item assessing whether being a San Francisco Giants fan is a reflection of who the participant is, the overall mean was 3.47

(SE = 0.28). The participants varied considerably in their identification with the Giants, but overall could not be considered highly identified Giants fans.

The three conditions differed in their self-integrity, $F(2, 88) = 3.39$, $p = .02$. We examined the condition differences using least significant difference (LSD) tests. Participants in the self-affirmation condition scored higher on the Self-Integrity Scale ($M = 6.17$, $SE = 0.13$) than did those in the no-affirmation condition ($M = 5.69$, $SE = 0.14$, $p = .012$). Participants in the awareness + affirmation condition fell between the other two conditions ($M = 5.85$, $SE = 0.19$) and were not significantly different from either the self-affirmation condition ($p = .17$) or the no-affirmation condition ($p = .50$); if anything, the awareness condition was closer to the no-affirmation control condition.

Effect of Affirmation Condition on Openness to the Report

We first analyzed responses to the four-item index assessing openness to the report (validity of essay, objectivity, intelligence of author, interest in reading *Game of Shadows*) with a one-way ANOVA and tested simple effects using LSD. Overall, there was an effect of the manipulation that approached significance, $F(2, 88) = 2.92$, $p = .059$ (see Figure 3). Those in the self-affirmation condition were more favorable toward the essay and its author ($M = 5.12$, $SE = 0.27$) than those in the no-affirmation condition ($M = 4.21$, $SE = 0.29$, $p = .018$). Those in the awareness + affirmation condition fell between the two groups ($M = 4.75$, $SE = 0.28$) and did not differ from those either in the self-affirmation condition ($p = .44$) or the no-affirmation condition ($p = .26$). However, the main effect of affirmation, as described below, was driven by the highly identified Giants fans.

Effect of Affirmation on Identity-Based Bias

We hypothesized that engaging in a self-affirmation would attenuate the bias from identification but that increased awareness that the affirmation could boost self-esteem would reduce the impact of the affirmation. To test these hypotheses we conducted multiple regression analyses using planned contrasts that tested the interaction between identification and the different affirmation conditions. Following Aiken and West (1991), we created planned contrasts to examine whether the relationship between identification and evaluation of the article varied across the three different affirmation conditions. We treated the no-affirmation condition as the baseline and examined whether the self-affirmation condition and the awareness + affirmation conditions deviated from this baseline. Thus, we created two dummy variables, the first with the self-affirmation condition coded 1 and the other two conditions (the no-affirmation condition and the awareness + affirmation condition) coded 0, the second with the awareness + affirmation condition coded 1 and the other two conditions (the no-affirmation condition and the self-affirmation condition) coded 0. Because they are entered simultaneously in the multiple regression analysis, the first contrast tests whether the self-affirmation condition differs from the no-affirmation condition, and the second contrast examines whether the awareness + affirmation condition differs from the no-affirmation condition. We then regressed the four-item index assessing openness to the report on the affirmation versus no-affirmation contrast, the awareness + affirmation versus no-affirmation contrast, San Francisco Giants identification (mean centered), and the two interactions (calculated by multiplying the pretest identification by each contrast). These results are displayed in Table 2.

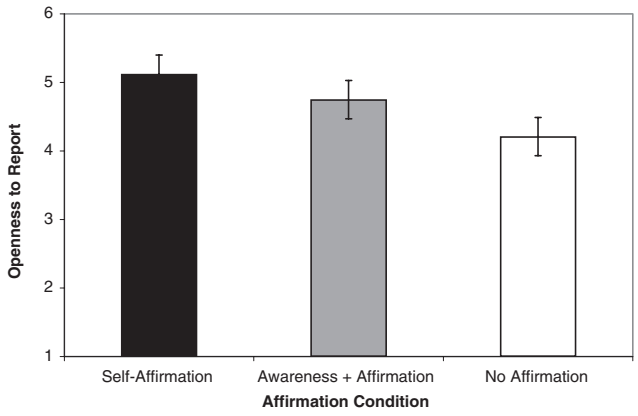

Figure 3. Openness to report as a function of affirmation status in Study 2. Error bars denote standard error.

The significant affirmation versus no-affirmation contrast indicates that, as noted above, those in the self-affirmation condition were more open to the article than those in the no-affirmation condition. The significant Affirmation × Identification interaction indicates that the relationship between identification and openness to the article differed between the affirmation and the no-affirmation conditions. The nonsignificant Awareness × Identification interaction indicates that the awareness + affirmation condition did not differ from the no-affirmation condition.

To examine the nature of these relationships, we conducted simple slopes analyses and estimated the values of the openness index at 1 *SD* above and below the mean on identification for the three conditions. In the no-affirmation condition, there was a strong negative relationship between identification and favorability ($\beta = -.39$, $p = .02$). Those who were highly identified Giants fans (estimated $M = 3.40$) were much less accepting of the information than those who were weakly identified Giants fans (estimated $M = 4.82$). In the self-affirmation condition, there was no relationship ($\beta = .06$, $p = .69$), as highly identified Giants fans (estimated $M = 5.21$) were as accepting of the information as weakly identified Giants fans (estimated $M = 5.01$). Thus, the self-affirmation enabled the highly identified Giants fans to be much more open to the critical information about Barry Bonds. Finally, participants in the awareness + affirmation condition ($\beta = -.18$, $p = .48$) evaluated the information in a similar manner as those in the no-affirmation condition, with those who were strongly identified Giants fans (estimated $M = 4.49$) trending toward being more critical than weakly identified Giants fans (estimated $M = 5.01$), although this negative relationship was not significant. Examining the estimated means of the highly identified Giants fans demonstrates the effect most clearly: In the no-affirmation condition, they were very critical of the article (estimated $M = 3.40$); in the self-affirmation condition, they were very open to the article (estimated $M = 5.21$); and in the awareness + affirmation condition, this openness was reduced (estimated $M = 4.49$). In sum, those in the no-affirmation condition exhibited

Table 2
Contrast Analysis Examining the Effects of San Francisco Giants Identification on Openness to Article Critical of Barry Bonds Across the Three Affirmation Conditions in Study 2

Predictor	*B*	*SE B*	*t*(85)	β
Identification	−.857	.343	−2.50*	−.420
Affirmation contrast	.999	.374	2.67**	.299
Awareness + affirmation contrast	.646	.473	1.37	.152
Identification × Affirmation Contrast	.980	.448	2.19*	.346
Identification × Awareness + Affirmation Contrast	.531	.701	0.76	.088

Note. The three affirmation conditions were no affirmation, self-affirmation, and awareness + affirmation. The affirmation contrast variable compares the self-affirmation condition to the no-affirmation condition, and the awareness + affirmation contrast variable compares the awareness + affirmation condition to the no-affirmation condition. Interactions involving these variables and identification indicate that the effect of identification in each condition differs from the no-affirmation condition.
* $p < .05$. ** $p < .01$.

strong identity-based bias, the self-affirmation eliminated this bias, and the instructions to heighten awareness attenuated the effect of the affirmation.

Assessments of Awareness

To examine participants' awareness of the manipulations, we first examined the open-ended responses to the question as to what factors may have influenced their responses to the essay about Barry Bonds. Two coders, who were unaware of condition, categorized the listed factors into one of five categories (which were derived from the participants' responses): information from the media (e.g., "facts I have read in the [San Jose] Mercury newspaper"), factors related to being a Giants/Barry Bonds fan (e.g., "Barry Bonds is one of the most driven people ever . . ."), knowledge of MLB policies and procedures (e.g., "Comparative info and stats on the baseball world of the past"), understanding of steroids and/or performance enhancing drugs (e.g., "I am a registered nurse and know about drugs"), and other comments (mostly editorial, e.g., "I'm not surprised or even interested in the transgressions of celebrities or sports figures, and it amazes me that anyone else is"). Overall, the two coders had 85.7% agreement on the categories, and a third coder resolved the differences. As in Study 1, no participants mentioned anything related to the self-affirmation activity (i.e., the first writing activity). In terms of the factors participants listed as impacting their assessments, 25% were related to the media, 16.7% were related to being a Giants/Bonds fan, 11.1% were related to MLB policy, 5.6% were related to knowledge about steroids, and 30.6% were other.

Next, we examined the ratings of the 11 potential influencing factors. As in Study 1, we standardized ratings of each of these factors within person and conducted a MANOVA with the 11 factors as the dependent variables and condition as the independent variable (see Figure 4). The multivariate analysis revealed no effects of condition, $F(22, 150) = 1.13$, $p = .32$. On the key awareness item, how much did "the first writing exercise" influence their evaluation of the article, there were no significant differences, $F(2, 84) = 1.16$, $p = .32$, between the self-affirmation

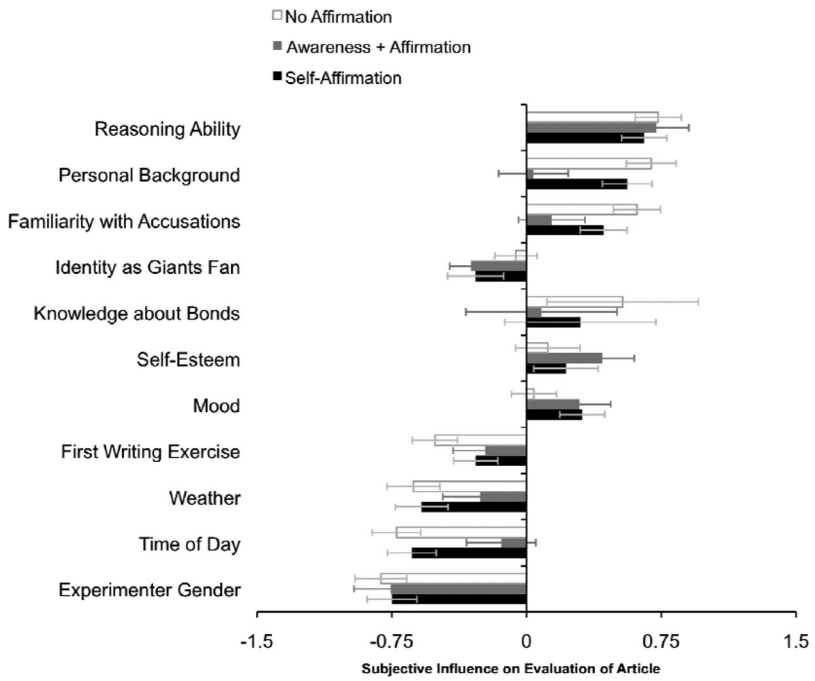

Figure 4. The subjective influence of 11 factors measured in Study 2. Higher numbers indicated that participants thought that factor was more influential. Error bars denote standard error. Note that "first writing exercise" referred to the affirmation manipulation.

condition ($M = -0.28$, $SE = 0.12$), awareness + affirmation condition ($M = -0.23$, $SE = 0.18$), and no-affirmation condition ($M = -0.51$, $SE = 0.13$).[5] Thus, although participants may have been aware that the self-affirmation affected their self-worth, as reflected in the increased scores on the Self-Integrity Scale in the affirmation condition, they did not indicate increased awareness that the affirmation activity affected their subsequent assessments.

In terms of general trends, participants across conditions thought that their evaluation of the Barry Bonds article was most influenced by their reasoning ability ($M = 0.69$) and least influenced by the gender of the experimenter ($M = -0.77$). They thought that the writing exercise, the self-affirmation manipulation, was the 8th most important factor out of the 11 ($M = -0.34$).

Awareness of Affirmation and Identity-Based Bias

As in Study 1, we examined the relationship between estimated influence of the affirmation and its actual influence. First, we examined whether awareness interacted with condition overall, and then we examined the three conditions separately. We used the same contrasts as in the identification analysis (comparing both the affirmation and the awareness + affirmation conditions to the no-affirmation condition) except we used the awareness of the impact of the writing exercise (the affirmation) as the predictor. There was no main effect of awareness ($\beta = -.043$, $p = .81$), nor was there an interaction between awareness and either contrast ($\beta = -.026$, $p = .87$, for affirmation contrast with no affirmation; $\beta = .044$, $p = .76$, for awareness contrast with no affirmation). However, this makes sense, as the real nature of the bias that was attenuated was observed not through the main effect of affirmation condition but in the interaction between identification and affirmation condition. That is, the important question is whether awareness interacts with identification in the affirmation condition: Does awareness of the impact of the affirmation undermine the effect of the affirmation on reducing the identity-based bias that was observed in the no-affirmation condition?

To examine this, we conducted separate analyses for each of the three conditions, where we regressed openness to the report on identification, the perceived influence of the manipulation, and the interaction (each mean centered). In the self-affirmation condition, there was a significant interaction ($\beta = -.561$, $p = .001$), but there was no interaction in the other two conditions (interaction $\beta = -.09$, $p = .69$, in the no-affirmation condition; interaction $\beta = .12$, $p = .77$, in the awareness + affirmation condition). We interpreted the interaction in the self-affirmation condition by estimating the predicted means at 1 *SD* above and below the mean on the two continuous predictors (identification and awareness). The affirmation was most effective among those individuals who reported that they were not influenced, as the highly identified Giants fans were more open (estimated $M = 7.07$) than the less identified Giants fans (estimated $M = 3.88$; $\beta = .66$, $p = .01$). By contrast, among those who said they were influenced within the affirmation condition, the highly identified Giants fans (estimated $M = 3.83$) were much less open to the article than the less identified fans (estimated $M = 6.73$; $\beta = -.87$, $p = .001$).[6] That is, their responses were very much like those in the no-affirmation condition, overall, with people responding with identity-defending bias. In sum, the predicted effects of the self-affirmation were observed only among those who said that they were not influenced by the manipulation; increased awareness was associated with decreased effectiveness of the affirmation.[7]

Discussion

Study 2 made three important demonstrations about awareness and self-affirmation. First, supportive of Hypothesis 2, participants were generally unaware of self-affirmation's influence. Participants did not spontaneously mention it as an influential factor and, when asked, thought the effect of the writing exercise on their evaluation of the Barry Bonds information would be rather small. Second, in partial support of Hypothesis 4, when participants were made aware of affirmation's influence (that it was designed to boost self-esteem), the effect on eliminating ideological bias was partially attenuated. Third, supportive of Hypothesis 3, the extent to which people thought that they were influenced was not positively correlated with the effectiveness of the manipulation. Indeed, in the self-affirmation condition, estimated influence was negatively correlated with the actual impact of the manipulation, as

[5] Two participants did not complete this measure, and 2 participants did not have any deviation (i.e., 1 circled all 5s and 1 circled all 9s). Because we standardized within person to demonstrate the relative assessment of impact of these items, we could not compute a standard deviation for these participants (and it did not appear that they took this measure seriously), and their data were thus omitted. Finally, 2 participants did not complete either one or three items of the awareness probe (importantly, they did complete the key item on perception of impact of the writing exercise, the affirmation manipulation). To keep these participants in the overall MANOVA, we assigned them the overall means for those particular items. This results in 84 degrees of freedom for the key univariate analyses following up the MANOVA. Further, examination of the individual items revealed a main effect on the item "personal beliefs," as those in the self-affirmation condition ($M = 0.60$, $SE = 0.14$) and the no-affirmation condition ($M = 0.69$, $SE = 0.14$) thought that their personal beliefs influenced their evaluation of the article more than did those in the awareness + affirmation condition ($M = 0.03$, $SE = 0.19$), $F(2, 82) = 4.02$, $p = .02$. There was also a main effect on the item "time of day," as those in the awareness + affirmation condition ($M = -0.14$, $SE = 0.19$) thought time of day influenced them more than did those in the self-affirmation condition ($M = -0.67$, $SE = 0.14$) or the no-affirmation condition ($M = -0.72$, $SE = 0.14$), $F(2, 82) = 3.39$, $p = .04$.

[6] It is interesting that awareness seemed to boost the effect of the affirmation on openness among weakly identified fans. One possible explanation derives from research on affirmation in nonthreatened domains (Briñol et al., 2007). This research finds that self-affirmations boost confidence in one's opinion, which can lead to greater openness to nonthreatening information. Considering that for weakly identified Giants fans, the anti-Barry Bonds article was value concordant, it is likely that the affirmation boosted their confidence in their opinion and that awareness did not attenuate this effect. Although the central concerns of this article are how affirmation affects people under threat, this finding raises interesting issues for future research.

[7] The preceding examination of awareness differs somewhat from that in Study 1 in that Study 1 revealed a significant Awareness × Condition interaction, whereas Study 2 revealed a significant Identification × Awareness interaction within the self-affirmation condition. This difference may stem from the overall level of math identification in Study 1 ($M = 6.18$ on the item assessing the importance of doing well on standardized tests) being much higher than the level of San Francisco Giants identification in Study 2 ($M = 4.22$ for the item assessing personal importance of being a San Francisco Giants fan, with both measures on 9-point scales).

only those participants who believed that they were not influenced by the affirmation exhibited the predicted identity-bias reduction effects.

Taken together, the first two studies suggest that self-affirmations are most effective when participants are unaware of their impact. The posttest awareness probe also yielded some illuminating findings. In general, people were unaware of the standard affirmation procedure, but when they did make a connection between the affirmation and the evaluation of the threatening information, the theorized affirmation effect was eliminated. However, it is important to recognize the inferential limits of these correlational analyses. To fully examine the effect of the awareness of the connection between the affirmation and the threatening information, we need to manipulate awareness of the link between the self-affirming task and subsequent judgments, an issue we turn to in Study 3.

Study 3

In Study 3, we experimentally manipulated the explicit connection between the threat and the affirmation. The first two studies indicated that affirmations are more effective (at improving performance and reducing identity-based bias, respectively) when participants believed that they were not influenced by the affirmation activity. When participants reported that they were influenced by the manipulation, the manipulation was ineffective. When participants were told that the writing activity could boost their self-esteem, the affirmation lost effectiveness relative to the standard affirmation condition (Study 2).

In Study 3, in addition to a standard self-affirmation condition (completing a values scale of a most important value) and an implicit self-affirmation condition (unscrambling sentences related to an important value), some participants were led to connect the standard affirmation with their responses along a measure of defensiveness. The latter condition will be termed an *explicit self-affirmation* because the connection between the affirmation and the threatening information was made explicit. Thus, in Study 3, we merely led some participants to suspect a link between the affirmation (the independent variable) and the measure of defensiveness (the dependent variable) without specifying the direction of the influence. Many self-affirmation studies (Cohen et al., 2000; Fein & Spencer, 1997; Sherman et al., 2000) use a "two studies" cover story precisely to eliminate this type of connection between the affirming activity and the subsequent dependent measure of interest. In an examination of Hypothesis 5, we tested whether recognition of such a link would undermine the affirmation's influence and exacerbate defensive biases because participants would become suspicious of the affirmation activity.

Previous research has suggested that when participants are aware of the potential influence of biasing contextual factors on their judgments, they adjust their judgments in the other direction (Bargh, 1992; Bargh et al., 1996; Stapel et al., 1998; Wegener & Petty, 1995). To use the terminology of Stapel et al. (1998), individuals are sensitive to the "smell of bias" and adjust their judgments to compensate for the effects of biasing stimuli in the direction of their naive theories (see also Wilson & Brekke, 1994). In Study 3, we examined whether enabling participants to potentially "smell" the biasing influence of the self-affirmation by making explicit the connection between the affirmation activity and the measure of defensiveness would invalidate the effectiveness of the affirmation.

In addition, we provided another test of Hypothesis 1, as we included an implicit affirmation manipulation in which participants unscrambled sentences designed to affirm an important value. Because Study 1 demonstrated an effect of an implicit affirmation (relative to a neutral control condition), we thought it would be instructive to compare an implicit affirmation (via sentence-unscrambling primes) with a standard affirmation in the present study.

Thus, Study 3 featured a continuum of awareness. At one end of the continuum was the implicit affirmation condition, where participants were aware of neither the value-relevance of the stimulus nor its potential impact. At the other end of the continuum was the explicit condition, where participants completed a standard self-affirmation (and hence were aware of the value-relevance of the stimulus) and were given instructions highlighting the link between the self-affirmation and the dependent variable. In the middle of this continuum was the standard self-affirmation condition, where participants were aware of value-relevance of the stimulus but unaware of its connection to the dependent variable—the state of affairs in a typical self-affirmation study. These three different conditions enable an examination of where the standard affirmation manipulation falls in terms of its impact on defensiveness as a function of awareness, relative to a fourth, no-affirmation condition.

Study 3 investigated the evaluation of comparative health risks. When asked to evaluate personal health risks, people think that they will experience better health outcomes than the average person (Weinstein, 1980; see also Chambers & Windschitl, 2004). This may be partially determined by the self-threatening nature of seeing the self as being at risk for disease, as previous research has shown that a self-affirmation can increase perceived risk for various diseases such as HIV (Sherman et al., 2000). Unlike in the first two studies, we presumed that health was relatively important to all participants, and thus we did not include identification as a moderator. The present study examines whether self-affirmation reduces this optimism bias and how awareness moderates that effect.

Method

Participants and Design

Participants were 83 introductory psychology students who were randomly assigned to one of four conditions: implicit affirmation, standard affirmation, explicit affirmation, or no affirmation. Data collection took place in two waves, first in fall 2003 ($n = 50$) and then in fall 2008 ($n = 33$); collection wave did not influence the results (i.e., wave was not a significant covariate and there was no interaction between wave and condition), so we combined the samples. The participants had a mean age of 18.44 years ($SD = 0.73$), consisted of 27 men and 55 women (1 did not report gender data), and included 72 European Americans, 4 Latinos, and 7 other/missing data.

Procedure

All participants completed a number of questionnaires in a study of "social judgment," but for explicit affirmation participants we altered the instructions to connect the affirmation to the dependent measure. In the implicit affirmation and standard affirmation conditions, participants were told that they would be completing two

separate studies, one on personal values and the other on evaluating health risk information. In the explicit affirmation condition, participants were told that the two elements were part of one study and were further told that "we are looking at the connection between peoples' values and their health beliefs," and in particular, how completing the values scale (the affirmation) may influence their health beliefs (the dependent measure). Finally, in the no-affirmation condition, participants were told that they would be completing two separate studies, but that the first one would be on evaluating health risk information. Thus, they completed the unrealistic optimism and no value-relevant manipulation.

As part of a pretest, all participants had rank ordered the importance of five values (political, social, aesthetic, religious, and theoretical), enabling the experimenter to prepare materials that ideographically emphasized the participant's most important value. Participants in the explicit and standard affirmation conditions completed a subscale from the Allport et al. (1960) study of values. The subscale listed a series of pairs of statements or response options, one of which corresponded to the participant's most important value, and the task was to circle which statement was more agreeable to the participant. This manipulation, used in prior self-affirmation research (e.g., Sherman et al., 2000; Tesser & Cornell, 1991), is theorized to affirm the self by focusing elaboration on a central value.

Implicit affirmation participants unscrambled sentences from the same values scale, with a cover story suggesting the study was a psycholinguistics investigation. For example, if religion was the most important value, the participant would be given, "it is important more for my child secure to more training religion in than athletics" and asked to unscramble it to something like, "It is more important for my child to secure training in religion than athletics." We made the unscrambling fairly easy so that participants could make the sentence in a way that would activate their most important value, and indeed participants unscrambled 82.7% (*Mdn* = 90%) of the sentences in a value-congruent manner.

After the affirmation manipulation, participants completed a "health beliefs" study, in which participants assessed their own health risks and the average student's health risks. Participants in the implicit and standard conditions were told that they had completed the first study, whereas participants in the explicit condition were reminded that "we're interested in the relationship between personal values and health beliefs."

All participants then completed the dependent measure, adapted from Weinstein and Klein (1995), in which they rated their personal health risks and the risks of the average student at their university for six health domains: heart disease, having a healthy child, skin cancer, living until age 85, exercising regularly, and developing a serious weight problem. For each of these health outcomes, participants made ratings on 10-point scales both for the self and for the average student. For example participants rated "What is the chance that you will develop a serious weight problem?" on a scale from 0 (*no chance to happen*) to 9 (*certain to happen*) and then rated "What is the chance that the average student—same sex as you—will develop a serious weight problem?" on an identical scale. All participants were then probed for suspicion and thoroughly debriefed.

Results

Analytic Strategy

We predicted that when participants were made aware of a link between the affirmation and the health beliefs questionnaire, they would exhibit the unrealistic optimism bias in their comparative health assessments in a manner akin to those who were not affirmed. In contrast, when participants were affirmed yet unaware, by either the standard affirmation or by the implicit affirmation, they would exhibit relatively less bias in their comparative health assessments. We conducted a planned contrast assigning weights of –1, –1, 1, 1 to the explicit, no-affirmation, standard, and implicit affirmation conditions, respectively, to test whether the explicit affirmation and the no-affirmation conditions differed from the implicit affirmation and the standard affirmation conditions. We tested two additional orthogonal contrasts assigning weights of 1, –1, 0, 0 and 0, 0, 1, –1 to the explicit, no-affirmation, standard, and implicit affirmation conditions, respectively. The first contrast tested whether there were differences between the explicit and no-affirmation conditions and the second contrast tested whether there were differences between the implicit and standard affirmation conditions.

Unrealistic Optimism

For each of the six measures we computed a difference score representing the degree to which the participant exhibited unrealistic optimism bias. That is, for negative items (e.g., heart disease) we subtracted the personal ratings from the average student ratings, and for positive items (e.g., having a healthy child) we subtracted the average student ratings from the personal ratings. Higher numbers represented more bias, and a score of 0 indicated that people thought themselves to be equally susceptible as average others. We averaged the unrealistic optimism bias for the six measures.[8]

The first planned contrast found that participants in both the implicit affirmation ($M = 0.46$, $SE = 0.22$) and the standard affirmation ($M = 0.49$, $SE = 0.21$) conditions showed less unrealistic optimism than did participants in the no-affirmation ($M = 0.83$, $SE = 0.18$)

[8] Data from 2 participants were discarded as outliers. One participant scored more than 3 *SD*s above the mean on the primary unrealistic optimism dependent variable. The second participant's score was an outlier on the negative health consequences items (i.e., a score of 7.75 when the mean was 1.67 and the standard deviation was 1.14; essentially the participant said he/she was at the lowest possible risk and that the average student was at the highest possible risk). We omitted these participants' data from further analyses. Also, the reliability of the six-item measure of unrealistic optimism was low ($\alpha = .46$). However, it is important to note that because people vary on their particular health risks (e.g., skin cancer vs. heart disease), the goal of the measure is to include multiple items so as to tap into an overall pattern of bias, although this heterogeneity may have reduced internal reliability. Given this relatively low alpha, we also conducted the analysis for a three-item measure (heart disease, exercise, and serious weight problem) that had higher reliability ($\alpha = .64$). The results were virtually identical to the reported analyses, as the contrast comparing implicit and standard versus explicit and no affirmation was significant, $t(79) = 2.23$, $p = .029$, with a similar pattern of means (implicit affirmation, $M = 0.67$, $SE = 0.42$; standard affirmation, $M = 1.23$, $SE = 0.35$; no affirmation, $M = 1.57$, $SE = 0.24$; and explicit affirmation, $M = 1.85$, $SE = 0.24$).

and explicit affirmation conditions ($M = 1.04$, $SE = 0.21$), $t(79) = 2.19$, $p = .031$ (see Figure 5). Neither additional orthogonal contrast was significant, $t(79) = 0.77$, $p = .45$, and $t(79) = 0.10$, $p = .92$, indicating that there was no difference between the implicit and standard affirmation conditions and no difference between the explicit and no-affirmation conditions, respectively. Affirmation, whether instantiated via implicit or standard means, reduced unrealistic optimism, but this effect was eliminated when participants were made aware of the link between the affirmation and the subsequent judgment task.

Discussion

Study 3 demonstrated that when a self-affirmation and the threatening information evaluation were explicitly linked, self-affirmation failed to reduce the bias, as those in the explicit affirmation condition were as biased as those who were not affirmed at all. In contrast, participants in the self-affirmation or implicit affirmation conditions had a reduced optimism bias relative to this explicit affirmation condition, seeing their risks as relatively more similar to the average student. That there was no difference between the implicit affirmation and standard affirmation conditions suggests that making the task more implicit may not increase the effectiveness of the affirmation. What appears to be critical is that participants did not see the affirmation as connected to the subsequent evaluation of the threatening information (cf. Bargh et al., 1996). When such a connection was explicitly made clear, the standard affirmation was rendered ineffective. Awareness of the process of self-affirmation—that it could influence subsequent self-relevant judgments—eliminated the effectiveness of the affirmation at reducing defensive judgments.

General Discussion

We examined five hypotheses relevant to the role of awareness in the self-affirmation process. First, we hypothesized that the self-affirmation process can occur even when the self-affirmation is delivered outside of deliberative awareness. Studies 1 and 3 found that participants who unscrambled value-relevant sentences demonstrated typical self-affirmation effects. In Study 1, the affirmation improved performance following threat, and in Study 3 the implicit affirmation reduced defensive bias. Notably, these studies suggest that implicit self-affirmations can be effective, but not, as in Study 3, that they are more effective than standard affirmations.

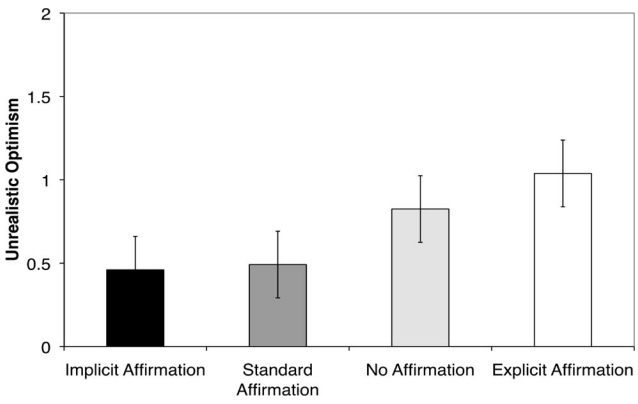

Figure 5. Unrealistic optimism in comparative health assessments as a function of affirmation condition in Study 3. Error bars denote standard error.

Second, we hypothesized that people are not aware that a self-affirmation manipulation will affect their behavior. In Studies 1 and 2, open-ended probes showed no evidence that people spontaneously attributed any influence to the manipulations, and when prompted to consider the effect of the affirmation manipulation (i.e., the "language task" in Study 1 or the "writing activity" in Study 2), they considered it to be a relatively unimportant influence on their behavior (akin to the time of day). Self-affirmation influenced behavior, but people were mostly unaware of it.

What happened as participants became more aware of the affirmation's influence? In Studies 1 and 2, and consistent with Hypothesis 3, it was only those participants who thought that they were not influenced by the affirmation whose behavior was actually influenced by the affirmation in the theorized direction. Increased estimated awareness was not positively related to increased effectiveness of the affirmation; if anything, increased affirmation awareness was associated with decreased affirmation effectiveness.

Finally, there was support for the hypotheses that affirmation effects would be eliminated when participants were told that the manipulation enhanced self-esteem (Study 2; Hypothesis 4) or that the manipulation was explicitly linked to the dependent measure (Study 3; Hypothesis 5). In Study 2, the awareness + affirmation condition was no different than the no-affirmation condition. In Study 3, participants made aware of the link between the affirmation and the dependent measure exhibited more bias than those who were affirmed, and the same amount of bias as those not affirmed. Experimentally induced awareness of the link between the affirmation and the subsequent evaluation of threatening information attenuated the beneficial effects of self-affirmation.

Why Does Awareness of Affirmation Attenuate the Effectiveness of Affirmation?

Given the consistency between the correlational and experimental findings demonstrating that increased affirmation awareness is associated with decreased affirmation effectiveness, the important theoretical question is why is this the case. Several different theoretical perspectives offer insight into this issue. First, people generally possess an illusion of objectivity, that is, they perceive themselves as interpreting information in a bottom-up, rational manner, and tend to see themselves as being less biased than the average person (Armor, 1999; Pronin et al., 2004; Ross & Ward, 1996). Moreover, people have naive theories about how they are influenced by the environment (Schwarz, 2004; Wegener & Petty, 1995), theories that in all likelihood do not include the potential that affirming the self in one domain will decrease threat responses in an alternative domain, the central finding of self-affirmation research (Sherman & Cohen, 2006; Steele, 1988). Consequently, when potential influences are made salient either by the experimental protocols or because the person just happens to suspect that he or she was influenced, the person is likely to try to correct for such influence. This result is consistent with the flexible correction

model of Wegener and Petty (1995). When people are made aware of a potential biasing influence, they attempt to correct for that bias. Moreover, people vary in their theories of bias, and these individual differences in awareness of influence predict the direction of people's corrections (Wegener & Petty, 1995).

This model could explain, in part, why the implicit self-affirming primes in Study 1 and the standard affirmation in Study 2 were generally ineffective when people suspected being influenced by them. Consistent with this notion, other research has found that when participants are aware of the potential influence of biasing contextual factors on their judgments, they adjust their judgments accordingly (Stapel et al., 1998). Bargh et al. (1996, p. 237) made a similar point, when they argued that in priming studies, it is "whether the individual is aware of the potential influence of the prime that is critical; diametrically opposite effects on judgments are obtained if the participant is aware versus not aware of a possible influence by the priming stimuli." Thus, people may realize that they are feeling more secure about who they are after completing the self-affirmations (and indeed, manipulation checks that used the Self-Integrity Scale in Studies 1 and 2 support this notion). Further, participants may realize that feeling better (e.g., more positive mood) could influence their judgments, and indeed the classic research by Schwarz and Clore (1983) demonstrates that alerting people to the weather affects their subsequent judgments and minimizes the actual effect of the weather. Our studies demonstrate this effect in the context of self-affirmation. Participants in Study 2 who believed that they were influenced by the standard self-affirmation manipulation exhibited opposite effects from those who believed that they were not influenced, and when we made the influence salient experimentally (in Study 2 and Study 3), the affirmation was no longer effective.

The second possibility, suggested by research on intrinsic motivation, is that more explicit attempts to affirm the self (of which people are probably more consciously aware) could end up being self-defeating, because by focusing on the end state people may undermine the affirmation process. For example, self-determination theory (Deci & Ryan, 1985) distinguishes among the reasons people strive for self-esteem (see also Sheldon, 2004). When self-esteem pursuit is done for calculated reasons, negative consequences are more likely, as it can undermine autonomy and higher order goals (Crocker & Park, 2004). As Schooler et al. (2003) put it, the pursuit of happiness can be self-defeating when conscious pursuit disrupts the happiness of hedonic pleasure.

Applying these notions to self-affirmation, the pursuit of self-affirmation may eliminate the beneficial effects of activities otherwise done for their intrinsic enjoyment. Self-affirmations typically secure the self in response to threat because they remind people of core values or other aspects of the self (e.g., religion or personal relationships) that will still be self-definitional regardless of the threat (e.g., academic failure). If individuals see the affirmational act as a means to an end (of feeling better about oneself), then it may lose that self-affirming quality of being a reminder of one's core values; the value may come to be seen as less important and intrinsically worthwhile. Indeed, when self-affirmations are more extrinsic (e.g., when people write about awards they could win as an artist), they lose any self-protective benefits (Schimel et al., 2004). Thus, the more explicit self-affirmation attempts are, and the more they lead individuals to focus on the benefits of the affirmation, the less effective the affirmations may be.

Furthermore, we speculate that if one is consciously affirming the self in response to threat in order to make oneself feel better, the affirmation may also be viewed as relevant to the threat itself. Given the research demonstrating that affirmations that are relevant to the threat backfire and even further threaten individuals (Blanton, Cooper, Skurnik, & Aronson, 1997; Sherman & Cohen, 2006), this is another possible reason why awareness may impede affirmation effectiveness. To close with an example: If, following a psychologically threatening event, a student just so happens to put on her college sweatshirt, perhaps she will experience some self-affirmation. On the other hand, perhaps if she consciously dons the sweatshirt to boost her spirits in response to the threat, she may find herself warmer, but no more affirmed.

Implications for Self-Affirmation Theory

The present studies make three important contributions to self-affirmation theory. First, they show that people are generally unaware of the self-affirmation process and that increased awareness of the process can impede the effectiveness of self-affirmation. Despite changes in performance (Study 1) and judgment (Study 2), people were generally unaware of the self-affirmation's influence. With heightened awareness—whether by measure (Studies 1 and 2) or manipulation (Studies 2 and 3)—the affirmations were generally less effective.

Second, implicit self-affirmations were demonstrated to be effective at reducing defensive responses to threat (Studies 1 and 3). It has now been documented in several studies that self-affirmations can influence implicit processes. For example, they can reduce the link between implicit self-judgments and implicit group judgments (Sherman & Kim, 2005), reduce race-based associations on the implicit associations test (Frantz, Cuddy, Burnett, Ray, & Hart, 2004), and influence implicit affect and rumination after failure (Koole, Smeets, van Knippenberg, & Dijksterhuis, 1999; see also Dijksterhuis, 2004). The present article is the first to show that self-affirmation can be instantiated and operated outside of deliberative awareness.

Third, our findings have implications for future field research. On the basis of the present results, we would like to offer a perhaps counterintuitive suggestion for those interested in applying self-affirmations in field settings. The key to an effective affirmation intervention may lie in the subtlety of its delivery and the minimalism of its administration. More transparent affirmations, those that are explicitly broadcast as "academic interventions" or "stress-reducers" by contrast, may raise awareness and reduce effectiveness. As Ross and Nisbett (1991) observed in their discussion of Kurt Lewin and the principle of situationism, "seemingly small situational factors that operate on important input or output channels will often exert gratifyingly large effects" (p. 11). It appears that self-affirmation can be one of these small situational factors. In the context of the present discussion, we would like to add that these seemingly small affirmation manipulations that secure self-integrity in otherwise threatening situations are most likely to be effective when participants are affirmed yet unaware.

References

Adams, G., Tormala, T. T., & O'Brien, L. T. (2006). The effect of self-affirmation on perception of racism. *Journal of Experimental Social Psychology, 42,* 616–626.

Aiken, L. S., & West, S. G. (1991). *Multiple regression: Testing and interpreting interactions.* Newbury Park, CA: Sage.

Allport, G. W., Vernon, P. E., & Lindzey, G. (1960). *Study of values* (3rd ed.). Boston: Houghton Mifflin.

Armor, D. A. (1999). The illusion of objectivity: A bias in the perception of freedom from bias. *Dissertation Abstracts International, 59*(B), 5163.

Aronson, J., Cohen, G. L., & Nail, P. R. (1999). Self-affirmation theory: An update and appraisal. In E. Harmon-Jones & J. Mills (Eds.), *Cognitive dissonance: Progress on a pivotal theory in social psychology* (pp. 127–148). Washington, DC: American Psychological Association.

Aronson, J., Lustina, M. J., Good, C., Keough, K., Steele, C. M., & Brown, J. (1999). When white men can't do math: Necessary and sufficient factors in stereotype threat. *Journal of Experimental Social Psychology, 35,* 29–46.

Bargh, J. A. (1990). Auto-motives: Preconscious determinants of social interaction. In E. T. Higgins & R. M. Sorrentino (Eds.), *Handbook of motivation and cognition* (Vol. 2, pp. 93–130). New York: Guilford Press.

Bargh, J. A. (1992). Does subliminality matter to social psychology? Awareness of the stimulus versus awareness of its influence. In R. F. Bornstein & T. S. Pittman (Eds.), *Perception without awareness: Cognitive, clinical, and social perspectives* (pp. 236–255). New York: Guilford Press.

Bargh, J. A., Chen, M., & Burrows, L. (1996). Automaticity of social behavior: Direct effects of trait construct and stereotype activation on action. *Journal of Personality and Social Psychology, 71,* 230–244.

Blanton, H., Cooper, J., Skurnik, I., & Aronson, J. (1997). When bad things happen to good feedback: Exacerbating the need for self-justification with self-affirmations. *Personality and Social Psychology Bulletin, 37,* 684–692.

Briñol, P., Petty, R. E., Gallardo, I., & DeMarree, K. G. (2007). The effect of self-affirmation in nonthreatening persuasion domains: Timing affects the process. *Personality and Social Psychology Bulletin, 33,* 1533–1546.

Chambers, J. R., & Windschitl, P. D. (2004). Biases in social comparative judgments: The role of nonmotivated factors in above-average and comparative-optimism effects. *Psychological Bulletin, 130,* 813–838.

Chartrand, T. L., & Bargh, J. A. (1996). Automatic activation of impression formation and memorization goals: Nonconscious goal priming reproduces effects of explicit task instructions. *Journal of Personality and Social Psychology, 71,* 464–478.

Cohen, G. L., Aronson, J., & Steele, C. M. (2000). When beliefs yield to evidence: Reducing biased evaluation by affirming the self. *Personality and Social Psychology Bulletin, 26,* 1151–1164.

Cohen, G. L., & Garcia, J. (2008). Identity, belonging, and achievement: A model, interventions, implications. *Current Directions in Psychological Science, 17,* 365–369.

Cohen, G. L., Garcia, J., Apfel, N., & Master, A. (2006, September 1). Reducing the racial achievement gap: A social-psychological intervention. *Science, 313,* 1307–1310.

Cohen, G. L., Garcia, J., Purdie-Vaughns, V., Apfel, N., & Brzustoski, P. (in press). Raising minority performance with a values-affirmation intervention: A two-year follow-up. *Science*.

Cohen, G. L., Garcia, J., & Sherman, D. K. (2009). [Development of a self-integrity scale]. Unpublished raw data.

Cohen, G. L., Sherman, D. K., Bastardi, A., McGoey, M., Hsu, A., & Ross, L. (2007). Bridging the partisan divide: Self-affirmation reduces ideological closed-mindedness and inflexibility in negotiation. *Journal of Personality and Social Psychology, 93,* 415–430.

Correll, J., Spencer, S. J., & Zanna, M. P. (2004). An affirmed self and an open mind: Self-affirmation and sensitivity to argument strength. *Journal of Experimental Social Psychology, 40,* 350–356.

Creswell, J. D., Welch, W., Taylor, S. E., Sherman, D. K., Gruenewald, T., & Mann, T. (2005). Affirmation of personal values buffers neuroendocrine and psychological stress responses. *Psychological Science, 16,* 846–851.

Crocker, J., Niiya, Y., & Mischkowski, D. (2008). Why does writing about important values reduce defensiveness? Self-affirmation and the role of positive, other-directed feelings. *Psychological Science, 19,* 740–747.

Crocker, J., & Park, L. E. (2004). The costly pursuit of self-esteem. *Psychological Bulletin, 130,* 392–414.

Davies, P. G., Spencer, S. J., Quinn, D. M., & Gerhardstein, R. (2002). Consuming images: How television commercials that elicit stereotype threat can restrain women academically and professionally. *Personality and Social Psychology Bulletin, 28,* 1615–1628.

Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior.* New York: Plenum Press.

Dijksterhuis, A. (2004). I like myself but I don't know why: Enhancing implicit self-esteem by subliminal evaluative conditioning. *Journal of Personality and Social Psychology, 86,* 345–355.

Ditto, P. H., & Lopez, D. F. (1992). Motivated skepticism: Use of differential decision criteria for preferred and nonpreferred conclusions. *Journal of Personality and Social Psychology, 63,* 569–584.

Epley, N., & Gilovich, T. (1999). Just going along: Nonconscious priming and conformity to social pressure. *Journal of Experimental Social Psychology, 35,* 578–589.

Fainaru-Wada, M., & Williams, L. (2006). *Game of shadows: Barry Bonds, BALCO, and the steroids scandal that rocked professional sports.* New York: Gotham.

Fein, S., & Spencer, S. J. (1997). Prejudice as self-image maintenance: Affirming the self through derogating others. *Journal of Personality and Social Psychology, 73,* 31–44.

Frantz, C. M., Cuddy, A. J. C., Burnett, M., Ray, A., & Hart, A. (2004). A threat in the computer: The race implicit association test as a stereotype threat experience. *Personality and Social Psychology Bulletin, 30,* 1611–1624.

Galinsky, A. D., Stone, J., & Cooper, J. (2000). The reinstatement of dissonance and psychological discomfort following failed affirmation. *European Journal of Social Psychology, 30,* 123–147.

Gilbert, D. T., Pinel, E. C., Wilson, T. D., Blumberg, S. J., & Wheatley, T. P. (1998). Immune neglect: A source of durability bias in affective forecasting. *Journal of Personality and Social Psychology, 75,* 617–638.

Gladwell, M. (2006). *Game of shadows.* Retrieved from http://gladwell.typepad.com/gladwellcom/2006/04/game_of_shadows.html

Glickman, R. (2002). *Optimal thinking: How to be your best self.* New York: Wiley.

Harris, P. R., Mayle, K., Mabbott, L., & Napper, L. (2007). Self-affirmation reduces smokers' defensiveness to graphic on-pack cigarette warning labels. *Health Psychology, 26,* 437–446.

Harris, P. R., & Napper, L. (2005). Self-affirmation and the biased processing of threatening health-risk information. *Personality and Social Psychology Bulletin, 31,* 1250–1263.

Hastorf, A., & Cantril, H. (1954). They saw a game: A case study. *Journal of Abnormal and Social Psychology, 49,* 129–134.

Jacks, J. Z., & O'Brien, M. E. (2004). Decreasing resistance by affirming the self. In E. S. Knowles & J. A. Linn (Eds.), *Resistance and persuasion* (pp. 235–257). Mahwah, NJ: Erlbaum.

Kay, A. C., & Ross, L. (2003). The perceptual push: The interplay of implicit cues and explicit situational construals on behavioral intentions in the prisoner's dilemma. *Journal of Experimental Social Psychology, 39,* 634–643.

Koole, S. L., Smeets, K., van Knippenberg, A., & Dijksterhuis, A. (1999). The cessation of rumination through self-affirmation. *Journal of Personality and Social Psychology, 77,* 111–125.

Kunda, Z. (1987). Motivated inference: Self-serving generation and evaluation of causal theories. *Journal of Personality and Social Psychology, 53,* 636–647.

Lau, R. R., & Russell, D. R. (1980). Attributions in the sports pages. *Journal of Personality and Social Psychology, 39,* 29–38.

Lord, C., Ross, L., & Lepper, M. R. (1979). Biased assimilation and attitude polarization: The effects of prior theories on subsequently considered evidence. *Journal of Personality and Social Psychology, 37,* 2098–2109.

Major, B., Spencer, S. J., Schmader, T., Wolfe, C., & Crocker, J. (1998). Coping with negative stereotypes about intellectual performance: The role of psychological disengagement. *Personality and Social Psychology Bulletin, 24,* 34–50.

Martens, A., Johns, M., Greenberg, J., & Schimel, J. (2006). Combating stereotype threat: The effect of self-affirmation on women's intellectual performance. *Journal of Experimental Social Psychology, 42,* 236–243.

Martin, L. L. (1986). Set/reset: Use and disuse of concepts in impression formation. *Journal of Personality and Social Psychology, 51,* 493–504.

McQueen, A., & Klein, W. (2006). Experimental manipulations of self-affirmation: A systematic review. *Self and Identity, 5,* 289–354.

Miller, D. T., & Ross, M. (1975). Self-serving biases in the attribution of causality: Fact or fiction? *Psychological Bulletin, 82,* 213–225.

Mitchell, G. J. (2007). *Report to the commissioner of baseball of an independent investigation into the illegal use of steroids and other performance enhancing substances by players in Major League Baseball.* Available from http://mlb.mlb.com/mlb/news/mitchell/index.jsp

Nussbaum, A. D., & Steele, C. M. (2007). Situational disengagement and persistence in the face of adversity. *Journal of Experimental Social Psychology, 43,* 127–134.

O'Brien, L. T., & Crandall, C. S. (2003). Stereotype threat and arousal: Effects on women's math performance. *Personality and Social Psychology Bulletin, 29,* 782–789.

Pronin, E., Gilovich, T., & Ross, L. (2004). Objectivity in the eye of the beholder: Divergent perceptions of bias in self versus others. *Psychological Review, 111,* 781–799.

Reed, M. B., & Aspinwall, L. G. (1998). Self-affirmation reduces biased processing of health-risk information. *Motivation and Emotion, 22,* 99–132.

Ross, L. D., & Nisbett, R. E. (1991). *The person and the situation: Perspectives of social psychology.* New York: McGraw Hill.

Ross, L., & Ward, A. (1996). Naive realism in everyday life: Implications for social conflict and misunderstanding. In E. S. Reed, E. Turiel, & T. Brown (Eds.), *Values and knowledge: The Jean Piaget symposium series* (pp. 103–135). Hillsdale, NJ: Erlbaum.

Schimel, J., Arndt, J., Banko, K. M., & Cook, A. (2004). Not all self-affirmations were created equal: The cognitive and social benefits of affirming the intrinsic (vs. extrinsic) self. *Social Cognition, 22,* 75–99.

Schmeichel, B. J., & Martens, A. (2005). Self-affirmation and mortality salience: Affirming values reduces worldview defense and death-thought accessibility. *Personality and Social Psychology Bulletin, 31,* 658–667.

Schooler, J. W., Ariely, D., & Loewenstein, G. (2003). The pursuit and monitoring of happiness can be self-defeating. In J. Carrillo & I. Brocas (Eds.), *Psychology and economics* (pp. 41–70). Oxford, United Kingdom: Oxford University Press.

Schwarz, N. (2004). Metacognitive experiences in consumer judgment and decision making. *Journal of Consumer Psychology, 14,* 332–348.

Schwarz, N., & Clore, G. L. (1983). Mood, misattribution, and judgments of well-being: Informative and directive functions of affective states. *Journal of Personality and Social Psychology, 45,* 513–523.

Sheldon, K. M. (2004). The benefits of a "sidelong" approach to self-esteem need-satisfaction: Comment on Crocker and Park (2004). *Psychological Bulletin, 130,* 421–424.

Sherman, D. K., Bunyan, D. P., Creswell, J. D., & Jaremka, L. M. (in press). Psychological vulnerability and stress: The effects of self-affirmation on sympathetic nervous system responses to naturalistic stressors. *Health Psychology.*

Sherman, D. K., & Cohen, G. L. (2002). Accepting threatening information: Self- affirmation and the reduction of defensive biases. *Current Directions in Psychological Science, 11,* 119–123.

Sherman, D. K., & Cohen, G. L. (2006). The psychology of self-defense: Self-affirmation theory. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 38, pp. 183–242). San Diego, CA: Academic Press.

Sherman, D. K., & Kim, H. S. (2005). Is there an "I" in "team"? The role of the self in group-serving judgments. *Journal of Personality and Social Psychology, 88,* 108–120.

Sherman, D. K., Kinias, Z., Major, B., Kim, H. S., & Prenovost, M. A. (2007). The group as a resource: Reducing biased attributions for group success and failure via group-affirmation. *Personality and Social Psychology Bulletin, 33,* 1100–1112.

Sherman, D. A. K., Nelson, L. D., & Steele, C. M. (2000). Do messages about health risks threaten the self? Increasing the acceptance of threatening health messages via self-affirmation. *Personality and Social Psychology Bulletin, 9,* 1046–1058.

Spencer, S. J., Steele, C. M., & Quinn, D. M. (1999). Stereotype threat and women's math performance. *Journal of Experimental Social Psychology, 35,* 4–28.

Stapel, D. A., Martin, L. L., & Schwarz, N. (1998). The smell of bias: What instigates correction processes in social judgments. *Personality and Social Psychology Bulletin, 24,* 797–806.

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: Academic Press.

Steele, C. M. (1997). A threat in the air: How stereotypes shape intellectual identity and performance. *American Psychologist, 52,* 613–629.

Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology, 69,* 797–811.

Steele, C. M., Spencer, S. J., & Aronson, J. (2002). Contending with group image: The psychology of stereotype and social identity threat. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 34, pp. 379–440). San Diego, CA: Academic Press.

Tesser, A., & Cornell, D. P. (1991). On the confluence of self-processes. *Journal of Experimental Social Psychology, 27,* 501–526.

Tesser, A., Crepaz, N., Collins, J. C., Cornell, D., & Beach, S. R. H. (2000). Confluence of self-esteem regulation mechanisms: On integrating the self-zoo. *Personality and Social Psychology Bulletin, 26,* 1476–1489.

Wegener, D. T., & Petty, R. E. (1995). Flexible correction processes in social judgment: The role of naive theories in corrections for perceived bias. *Journal of Personality and Social Psychology, 68,* 36–51.

Weinstein, N. D. (1980). Unrealistic optimism about future life events. *Journal of Personality and Social Psychology, 39,* 806–820.

Weinstein, N. D., & Klein, W. M. (1995). Resistance of personal risk perceptions to debiasing interventions. *Health Psychology, 14,* 132–140.

Wilson, T. D. (2002). *Strangers to ourselves: Discovering the adaptive unconscious.* Cambridge, MA: Belknap Press.

Wilson, T. D., & Brekke, N. (1994). Mental contamination and mental correction. *Psychological Bulletin, 116,* 117–142.

(*Appendix follows*)

Appendix

Self-Integrity Scale

Please indicate your agreement with the statements below by writing the appropriate number next to the statement using the following scale.

Strongly disagree		Disagree		Agree		Strongly agree
1	2	3	4	5	6	7

_____ 1. I have the ability and skills to deal with whatever comes my way.

_____ 2. I feel that I'm basically a moral person.

_____ 3. On the whole, I am a capable person.

_____ 4. I am a good person.

_____ 5. When I think about the future, I'm confident that I can meet the challenges that I will face.

_____ 6. I try to do the right thing.

_____ 7. Even though there is always room for self-improvement, I feel a sense of completeness about who I fundamentally am.

_____ 8. I am comfortable with who I am.

Received May 7, 2008
Revision received December 20, 2008
Accepted January 15, 2009 ■