Article

Self-Affirmations Provide a Broader Perspective on Self-Threat

Personality and Social Psychology Bulletin
2015, Vol. 41(1) 3–18

DOI: 10.1177/0146167214554956
pspb.sagepub.com
SAGE

Clayton R. Critcher[1] and David Dunning[2]

Abstract

We present an "affirmation as perspective" model of how self-affirmations alleviate threat and defensiveness. Self-threats dominate the working self-concept, leading to a constricted self disproportionately influenced by the threat. Self-affirmations expand the size of the working self-concept, offering a broader perspective in which the threat appears more narrow and self-worth realigns with broader dispositional self-views (Experiment 1). Self-affirmed participants, relative to those not affirmed, indicated that threatened self-aspects were less all-defining of the self (although just as important), and this broader perspective on the threat mediated self-affirmation's reduction of defensiveness (Experiment 2). Finally, having participants complete a simple perspective exercise, which offered a broader perspective on the self without prompting affirmational thinking (Experiment 3a), reduced defensiveness in a manner equivalent to and redundant with a standard self-affirmation manipulation (Experiment 3b). The present model offers a unifying account for a wide variety of seemingly unrelated findings and mysteries in the self-affirmation literature.

Keywords

self-affirmation, defensiveness, threat, perspective, trivialization

Received May 7, 2013; revision accepted September 7, 2014

People wish to view themselves as competent, compassionate, and worthy individuals, but their attitudes, intentions, and behaviors do not always match this ideal. As such, people possess an eclectic toolkit of strategies for defending their positive self-views even when their thoughts, their behavior, or external events call those self-views into question. People might deal with a threat to self-esteem directly by minimizing or dismissing it (Kunda, 1990), but people also address threats indirectly by bolstering some other aspect of their self-concept (e.g., Dunning, Leuenberger, & Sherman, 1995; Steele, 1975).

This last observation lies at the heart of self-affirmation theory (Sherman & Cohen, 2006; Steele, 1988), which suggests that bolstering the self in one important domain buffers the impact of threats in another. For example, although heavy alcohol users have a tendency to downplay their health risks, they are less defensive toward alcohol-related risks after writing about an important value bearing no connection whatsoever to alcohol consumption (Harris & Napper, 2005). Self-affirmations have been shown to forestall a variety of defensive responses (see Sherman & Cohen, 2006, for a review), such as outgroup derogation (Fein & Spencer, 1997; Stone, Whitehead, Schmader, & Focella, 2011), closed-minded negotiation (Cohen et al., 2007), and the stubborn persistence of unhealthy habits (Epton & Harris, 2008). Furthermore, affirmations alleviate the experience of threat itself. They reduce health-deteriorating stress (Sherman, Bunyan, Creswell, & Jaremka, 2009) and performance-inhibiting stereotype threat (Cohen, Garcia, Apfel, & Master, 2006; Martens, Johns, Greenberg, & Schimel, 2006).

What has been less clearly documented is *why* self-affirmation works. A typical way of representing this question is illustrated in Figure 1. By this account, the self occasionally encounters a threat that damages the worth or integrity of the self as a whole (Figure 1a). As Sherman and Cohen (2006) explained, "The self-system is activated when a person experiences a threat to an important self-conception or image . . . these events are threatening because they have implications for a person's overall sense of self-integrity" (pp. 187-188). In response, affirmation purportedly adds a sense of self-worth, thereby restoring integrity to the self (Figure 1b). Sherman and Cohen (2006) noted that "although a defensive bias can restore self-integrity" (p. 186), so can an affirmation, which itself is "an opportunity to restore [one's] self-integrity" (p. 201).

[1]University of California, Berkeley, USA
[2]Cornell University, Ithaca, NY, USA

Corresponding Author:
Clayton R. Critcher, University of California, Berkeley, Haas School of Busness, 545 Student Services Building, #1900, Berkeley, CA 94720-1900, USA.
Email: ClaytonCritcher@haas.berkeley.edu

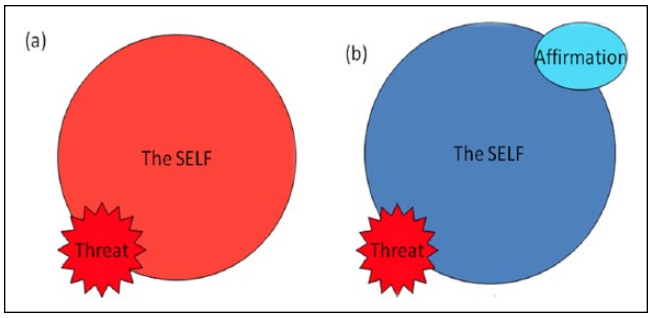

Figure 1. A typical conceptualization of self-affirmations' effects: (a) A threat damages the integrity of the self, and (b) an affirmation heals the self. From this perspective, answering how self-affirmations restore self-integrity is a question of what key "booster" it is that affirmations provide "The Self."

As such, affirmation is portrayed as an "integrity" booster shot, raising a person's sense that he or she is a coherent, moral, and adequate individual. In describing why threat would inspire either defensiveness or a need to self-affirm, Sherman and Hartson (2011) noted, "When the goal of global self-integrity maintenance is threatened, people seek means of re-affirming their self-integrity" (p. 130). From this conception, it is natural to ask what active ingredient this booster shot contains that augments integrity and thus eliminates defensiveness. The challenge then is to identify what self-affirmations add to the self that then mediates the reduction in defensiveness. Despite some scattered meditational support, no such *consistent* mediator has been identified (see Crocker, Niiya, & Mischkowski, 2008; Sherman & Cohen, 2006).

We suggest that this typical way of approaching self-affirmation research, although plausible, may be misleading in three ways. First, it inappropriately treats the self as a singular or invariant construct. Second, the question focuses narrowly on what affirmation does, rather than on what the threat does that self-affirmation interrupts. Third, a problem emerges when the question of how self-affirmations work is reduced to the question of what mediates the link between self-affirmation and defensiveness. As an example, Taylor and Walton (2011) found that self-affirmation reduces the performance-debilitating impact of stereotype threat and that this effect was mediated through stereotype suppression. However, it would be improper to conclude that self-affirmations reduced threat by reducing stereotype suppression (and the authors did not draw such a conclusion). Instead, the purported mediator may tell more about what explains a threat-defensiveness response (in this case, that stereotype suppression led to stereotype-threat-induced performance decrements) instead of how that threat response is eliminated by self-affirmation.

Given this, we suggest that understanding the impact of self-affirmation requires a more diverse approach than merely searching for a consistent mediator. Instead, any such empirical approach should help elucidate how affirmations undo self-threat. This approach places the focus on what threat does and how affirmation may undo it.

Affirmations Provide Perspective

We propose that self-affirmation acts as a buffer against threat because both self-affirmations and threats, alike, alter the very nature of the "self" being threatened. Like many before us, we posit that the representations and identities that compose one's self-concept in consciousness vary from moment to moment (Figure 2a). These salient contents compose the working self-concept (Markus & Wurf, 1987). In the face of failure or threat, the damaged identity may dominate one's working self-concept, flooding it with negative cognitions and emotions that implicate one's sense of self-worth (e.g., Koole, Smeets, van Knippenberg, & Dijksterhuis, 1999). In essence, the contents of the working self-concept are narrowed (Figure 2b), leading the threatened domain to loom large. Because of this, feelings and beliefs attached to this specific self-facet are disproportionately weighted in one's momentary (threatened) self-evaluation.

We suggest that affirmations blunt the impact of the threat because they expand the contents of the working self-concept—thus narrowing the scope of any threat (Figure 2c). This broadened perspective reminds people that the threatened domain is not *all* that defines the self, and so it mitigates the evaluative implications that a threat to any single identity has on perceptions of the self as a whole. The threat becomes not about *the self* but only about one narrowed aspect of the self. As a result, ego repair is less necessary.

Our approach shares similarities with logic relied upon in the self-complexity literature, which has studied why some people, *dispositionally*, weather threats better than do others (Linville, 1985, 1987). Specifically, research on self-complexity suggests that selves differ in terms of how many identities they have and how differentiated those identities are. Because threat will implicate a larger portion of the overall self of low-complex (vs. high-complex) individuals, such people show decreased mood (Linville, 1985; Niedenthal, Setterlund, & Wherry, 1992) and defensiveness (Dixon & Baumeister, 1991) in the face of negative feedback or threat.

We propose that under threat, the constricted nature of the working self-concept means it takes on a property of low-complex selves—the threat now implicates a larger portion of the (active) self-concept. As a result, threats loom subjectively larger and the negativity associated with that identity is an evaluation that dominates one's sense of self. With fewer other identities active to help dampen the evaluative impact of the threat, the stage is set for depressed self-integrity and defensiveness.

Although the idea that self-affirmation "expands" or offers perspective on the self—and, thus, makes any threat seem more narrow—has not been studied explicitly, work by Simon, Greenberg, and Brehm (1995, Study 3) indirectly

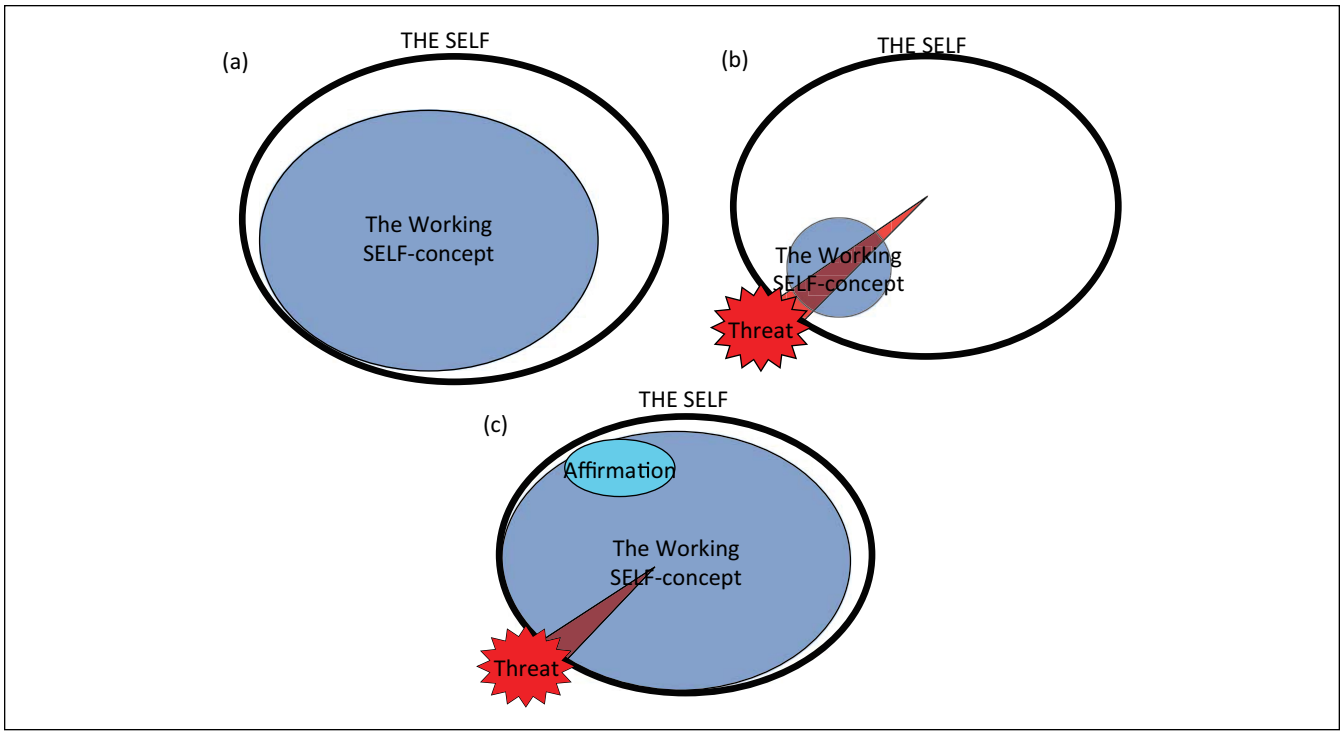

Figure 2. How threat and affirmation unfold according to the affirmation as perspective model.
Note. The model (a) recognizes that at any given time, only a subset of one's self—one's *working self-concept*—is accessible. (b) When threatened, the working self-concept constricts and ruminates on the threatened identity. As such, one's sense of self-worth is narrowly tied to the threatened domain, which dominates one's self-definition. (c) An affirmation broadens the working self-concept, thereby providing perspective on the threat, which should no longer loom as large in the self-concept or dominate one's feelings of self-worth.

suggests that affirmations may offer perspective by rendering threats a narrower part of oneself. They tested whether self-affirmations led to the trivialization of threat. Unexpectedly, their four trivialization items loaded on two distinct factors. Only one of those factors—reflected by a single item—was influenced by self-affirmation. Affirmed participants indicated that "in the grand scheme of things," the threatening behavior they had just engaged in was not all that significant. In our view, this item may not so much have assessed trivialization as an expanded perspective on the threat. The item asked not simply about the importance of the act itself but also about its relative significance in the context of the broader self. The items in the second, orthogonal factor (unaffected by self-affirmation) focused directly on the importance of threat or the threatened domain.

Although our perspective account has not been previously articulated or directly tested, there are several additional programs of research that make our reasoning plausible. First, evidence suggests that threat *constricts* attention and thought to focus on the source of threat. For example, dot probe paradigms have shown that attention is immediately drawn to a threatening picture over a neutral one (Mogg & Bradley, 1999). Although these findings were initially offered as evidence of attention capture by threatening stimuli (e.g., Mathews, Mackintosh, & Fulcher, 1997), more refined techniques showed that these effects are due to the difficulty of disengaging from threatening information rather than attention capture (Koster, Crombez, Verschuere, & De Houwer, 2004). Note that the difficulty of disengagement—applied to threats detected in one's self instead of one's world—is most consistent with our account. Threatened aspects of the self may engulf one's self-perception and keep it engulfed due to a difficulty of disengaging from the threat.

That a threatened identity may dominate the working self-concept until an outside intervention assists with disengagement was demonstrated in a recent investigation of stereotype threat. Rydell, McConnell, and Beilock (2009) gave women a task on which they feared they might confirm the cultural stereotype that women are bad at math. In this circumstance, their (threatened) female identity loomed large in their working self-concepts. However, when reminded of their identity as a college student—an identity associated with competence at math—the salience of their gender identity faded. That is, a threatened identity loomed large in the self-concept until another, positive identity was presented.

Although using similar language, recent research by Wakslak and Trope (2009) has a different take on self-affirmation. They emphasize that because self-affirmations focus people on abstract qualities of the self, people are

procedurally primed to view all stimuli in more abstract terms. Thus, even though they note that affirmations can help one "focus on the big picture" (p. 927), their procedural priming account would predict that affirmations lead people to see all stimuli, including threats, in more abstract terms. It is worth noting that even though our accounts discuss different consequences of self-affirmation, the two lines of research have very different aims, so they are not in competition.

Overview of the Present Experiments

We present four studies that, in combination, test whether self-affirmations enhance perspective on the self, and whether this perspective-granting property explains how self-affirmations reduce threat and defensiveness. Experiment 1 tested whether self-affirmations counteract threat by providing a broader perspective on the self. If people's momentary sense of self-worth is determined by what identities are active in one's working self-concept, and if threats constrict the working self-concept, then momentary feelings of self-worth during threat should be narrowly tethered to feelings about the threatened domain, and thus depressed (see Figure 2b). However, if affirmations buffer threat by expanding the self, then threatened participants' feelings of self-worth should remain high and align instead with one's more global, dispositional self-evaluations (see Figure 2c).

Experiment 2 tested whether self-affirmation changed participants' phenomenological sense of perspective: We hypothesized that affirmed participants (compared with the non-affirmed) would see the threat as just one part of a larger self, and thus not something that posed a broader threat. Crucially, we note (and test) that this narrowing of the scope is different from *trivializing* that aspect of self—that is, merely seeing it as less important (Correll, Spencer, & Zanna, 2004; Koole et al., 1999). Furthermore, we expected that perspective (but not trivialization) would mediate the effect of self-affirmation on defensiveness reduction.

Finally, we developed a new perspective exercise that retained some aspects of a standard self-affirmation (i.e., the ability to offer perspective), but that eliminated other aspects (the instigation of affirmational thinking, the writing task, the selective focus on an important identity). After validating this new manipulation (Experiment 3a), we tested whether it reduced defensiveness in a manner that is equivalent to (and redundant with) a standard self-affirmation intervention (Experiment 3b).

For each study, we did not determine a specific sample size in advance, but instead collected data until a stopping point naturally associated with the academic calendar (Fall Break, Thanksgiving Break, Spring Break, or the end of the semester) or, in one instance, the due date for the revision of this manuscript. These stopping dates permitted us to achieve good statistical power through reasonably large sample sizes. Across our five studies (pilot study, Experiments 1-3b), we averaged 68.6 participants per condition, or 67.8 per condition after exclusions (described below).

Pilot Study

In presenting our hypotheses, we have assumed that when threatened identities are salient in the working self-concept, people experience threat, feel their self-worth is tied to this salient identity, and even have a phenomenological sense of diminished perspective. Before turning to a direct examination of self-affirmation, we conducted a pilot study to test whether the accessibility of a threatened identity in the working self-concept has these properties. If so, it would provide a reasonable foundation from which our main experiments could build.

We showed 108 undergraduates at the University of California, Berkeley, a video that described an actually fictitious medical condition (TAA deficiency) that supposedly afflicted 20% of the population. Following the video, participants completed a "TAA Deficiency Lifestyle Questionnaire" that included 25 questions. Finally, participants were told that they could learn their personalized risk assessment for TAA Deficiency at the study's end. Howell and Shepperd (2012) used this procedure to successfully induce threat. We hypothesized that the salience of the threatened (health) identity in the working self-concept would predict greater threat and reduced perspective.

First, we measured the degree to which participants' health identity was active using a single-category Implicit Association Test (IAT; Karpinski & Steinman, 2006). This permitted us to compare how easily participants associated words related to the self (*I, I'm me, mine, my, myself, self*) to words related to health (*diagnosis, diet, exercise, health, wellness*), or a comparison category relevant to all of our participants, being a student (*academics, grades, pupil, school, student*).

We then collected our measures of threat and perspective. We assessed self-threat using another single-category IAT. This assessed the ease of associating words related to the self with words related to threat (*afraid, distressed, nervous, scared, uneasy*) versus security (*confident, determined, good, proud, strong*). Next, we measured *perspective* by having participants rate from 1 (*not at all*) to 9 (*very much so*), "If I learned I had TAA Deficiency, this deficiency would be specific enough that other aspects of myself would make up for it." We expected the accessibility of the threatened identity in the self to negatively relate to this self-reported *perspective*, but be unrelated to a measure of *trivialization* ("It is not that important to me whether I have TAA deficiency."). Finally, we had participants imagine both that they learned that they *did* have TAA Deficiency, and also to imagine that they *did not* have TAA Deficiency. Participants forecasted how they would feel with 16 measures—8 assessing positive feelings of self-worth (e.g., pleased with self) and 8 items from the Positive and Negative Affect Schedule (PANAS; for example, scared; Watson, Clark, & Tellegen, 1988). We

reverse-scored negative items, so that higher scores would reflect a greater sense that one would feel positively in response to the news. We then subtracted anticipated feelings of learning one did have the disease from anticipated feelings if one learned one did not have the disease. This difference reflects *anticipated perspective*, an estimate that one's feelings of self-worth will be narrowly contingent on the presence or absence of the threat.

Consistent with our predictions, the more participants' health identity loomed large in the working self-concept, they showed greater self-threat, $r(106) = .31$, $p = .001$, reported less perspective, $r(106) = -.19$, $p = .05$—but no more trivialization, $r(106) = -.07$, *ns*—and had diminished anticipated perspective, $r(106) = -.19$, $p = .05$. As our account would predict, when threats loom large in the self, they are related to threat, diminished perspective, and (anticipated) feelings of self-worth that are more contingent on that salient identity. These findings are also helpful in validating that our measures of perspective (which, in adapted form, are used in Experiments 1 and 2) relate to the degree to which threats loom large in the working self-concept.

Experiment 1

Affirmations are said to restore a sense of worth to the self (Steele, 1988). Sherman and Hartson (2011) noted that affirmations provide a psychological buffer for individuals to accept threatening information without sacrificing self-perceptions as globally competent and worthy individuals. By our account, threats narrow one's perspective, such that the threat and the threatened domain loom large and determine one's depressed sense of self-worth. Affirmations restore a broader perspective on the self, blunting the impact of a narrow threat, thereby permitting the self to draw on its broader dispositional resources to determine elevated self-worth.

To test these assertions, we threatened participants by having them fail on an intellectual task. We then assessed their situational feelings of self-worth. We expected that non-affirmed participants would have depressed feelings of self-worth largely driven by their sense of competence in the threatened domain (i.e., narrow perspective). We predicted that the affirmed would instead have elevated feelings of self-worth decoupled from the threat and instead aligned with their broader dispositional self-evaluations (i.e., broadened perspective). This would be evidence that (a) a threatened identity dominates the working self-concept thereby providing a narrow perspective on the self with an accompanying depressed sense of worth,[1] but (b) this threat can be prevented by a self-affirmation that restores a broader perspective on the self.

Method

Participants and design. Eighty-two students at Cornell University were randomly assigned to an affirmation or control condition. Seven participants (3 control, 4 affirmation) were excluded from all analyses because during debriefing they indicated that they believed the test to be "rigged" or "intended to make them feel bad about themselves."

Procedure. When signing up for the study (at least 24 hr before the lab component), participants completed a measure of dispositional self-esteem, modified from Rosenberg (1965). This served as a *broad self-assessment*.

Once participants arrived at the lab, those assigned to the affirmation condition completed a values-based affirmation. Participants ranked eight values or skills (e.g., adventure in life, financial success).[2] Participants then took 3 min to write about why their most valued domain was meaningful in their own life.[3] We positioned the self-affirmation manipulation prior to the test so that it would not serve as a distraction in between the threat and the final measures, a worry sometimes expressed when self-affirmations are positioned later (Steele, 1988).

Next, all participants were told that they would complete a test that measured "integrative orientation ability," supposedly a test of "creative thinking skills that are particularly diagnostic of success in professional careers." The test was a modification of the Remote Associates Test (Mednick, 1962) used in previous research to induce threat (Critcher, Dunning, & Armor, 2010). Participants received 15 word triads. For each triad, participants had to generate a fourth word that connected to each of the provided words. For example, one triad read "STALK–TRAINER–KING."[4]

After the remote associates test, participants completed 14 items measuring their feelings of self-worth, asking them "how well each statement characterizes how you feel about yourself right now." A principal components analysis with varimax rotation found that items loaded on two orthogonal factors. As listed in Appendix A, 8 items loaded on the *positive feelings of self-worth* factor, while 6 items loaded on the *negative feelings of self-worth* factor. We standardized and averaged the items, reverse scoring where necessary, to create a positive feelings ($\alpha = .92$) and negative feelings composite ($\alpha = .91$).

To assess participants' views of their own abilities specific to the tested domain, participants indicated their agreement that "I feel I am pretty good at tests like the one I took today," on a scale from 1 (*not at all*) to 9 (*extremely*). This served as a *narrow self-assessment*.

Results

Confirming that the remote associates test was very difficult, participants correctly answered only 2.6 out of 15 ($SD = 2.01$) questions, on average. Performance did not differ between conditions, $t(73) = 1.46$, $p > .14$, $d = .34$, 95% confidence interval (CI) = [−0.25, 1.61]. However, affirmed participants had higher positive feelings of self-worth ($M = 0.16$, $SD = 0.74$) compared with those in the control

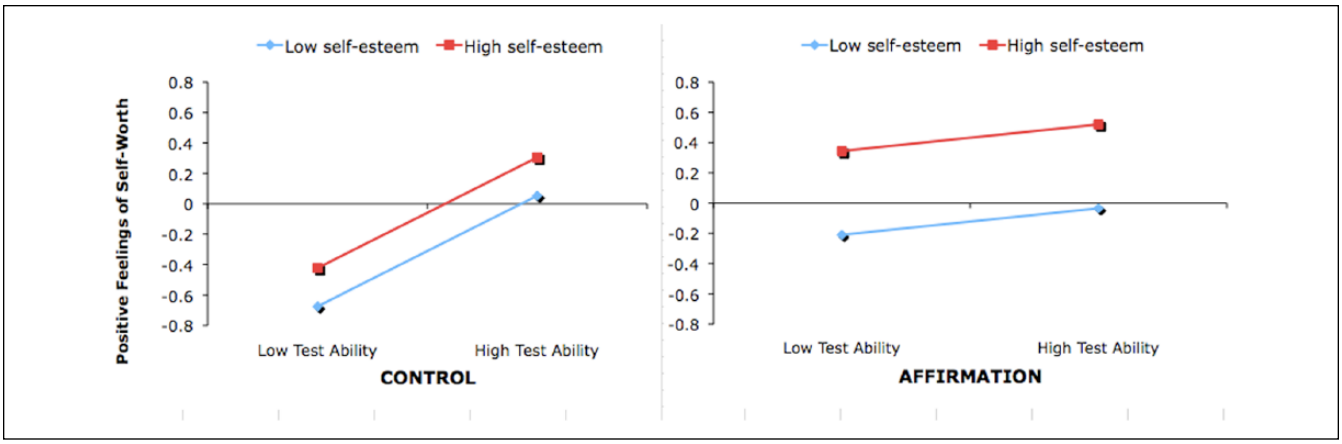

Figure 3. Positive feelings of self-worth, by condition, dispositional self-esteem, and self-perceived ability in the threatened domain (Experiment 1).

Note. Plotted values are the predicted values for a participant in the specified condition who is ± 1 *SD* in dispositional self-esteem (broad source of self-assessment) or self-perceived ability in the threatened domain (narrow source of self-assessment).

condition ($M = -0.21$, $SD = 0.86$), $t(73) = 1.95$, $p = .05$, $d = .46$, 95% CI = [−0.01, 0.73]. The groups did not differ in their negative feelings of self-worth, $t < 1$.

By our account, these differences in positive feelings of self-worth should be traceable to the differential influence of narrow versus broad sources of self-assessment in the affirmation versus control condition. First, we conducted simple slopes analyses to determine the predictive power of broad and narrow self-assessment on positive feelings of self-worth in affirmation versus control participants (Figure 3). Second, we performed a single statistical test that most directly assesses our perspective hypothesis—that is, that the relative predictive power of broad versus narrow self-assessment differs by affirmation condition.

Participants' broad and narrow self-assessments were uncorrelated, $r = -.10$, addressing any concern that the beta weights reported below were distorted by a multicollinearity problem. In the control condition, momentary positive feelings of self-worth were related to the narrow self-assessment in the threatened domain, $\beta = .59$, $t(64) = 3.07$, $p = .003$, 95% CI = [0.21, 0.98], but not to dispositional self-esteem, $\beta = .17$, $t < 1$. This reflects low perspective, in that one's sense of worth was dominated by the focal, but narrow threat. In contrast, affirmed participants' positive feelings of self-worth were not associated with their specific assessment of ability in the threatened domain, $\beta = .08$, $t < 1$, but were tied to broader dispositional self-esteem, $\beta = .42$, $t(64) = 3.22$, $p = .002$, 95% CI = [0.16, 0.68]. This reflects a broad perspective on the self, in that momentary feelings of self-worth were dissociated from the narrow threat and instead aligned with a broader dispositional self-evaluation.

Although this pattern is consistent with our hypotheses, we wanted to provide a single omnibus test of our predictions. Essentially, we wanted to test whether the relative contribution of the two predictors of self-worth in the control condition (narrow vs. broad: .59 vs. .17) differed from that in the self-affirmation condition (narrow vs. broad: .08 vs. 42). Readers can think of this as a proposed 2 × 2 interaction, but focusing on beta weights rather than mean responses. We regressed positive feelings of self-worth on affirmation condition (−1 = control, +1 = affirmation), narrow self-assessment (standardized), broad self-assessment (standardized), and a modified slate of the 3 two-way and 1 three-way interaction terms. That is, in place of the Affirmation × Narrow and Affirmation × Broad terms, we substituted in one predictor that was the sum of these two expressions, and one predictor that was the difference between these two terms. The difference score provides the crucial test of whether the relative predictive power of the narrow and broad self-assessments differ across conditions.[5] Confirming our central hypothesis that the degree of perspective differed by condition, the difference score was the only new term to reach significance, $\beta = .29$, $t(64) = 2.42$, $p = .02$, 95% CI = [0.05, 0.52].[6]

Discussion

In short, these data suggest that under threat, people have low perspective, with their sense of self-worth tethered to depressed feelings linked to the threatened domain. In other words, threats without affirmations have a large impact on people's sense of self-worth. Self-affirmations broadened perspective, pushing people's sense of self-worth to instead align with their broader self-views (i.e., dispositional self-esteem). This suggests that self-affirmations restore self-worth by "undoing" an otherwise constricted perspective under threat.

In our study, participants received honest feedback. This meant that there was variability in the actual number of items participants knew they answered correctly. Note that if some of the variability in participants' specific self-assessments comes

from the variation in their test performance (and as expected, they are correlated, $r = .35$, $p = .001$), then this is a source of *meaningful* variation that should not be controlled for. Nonetheless, all analyses and patterns of significance remain unchanged if test performance is included as a covariate.

Finally, we wish to stress the value of examining affirmations' effects on feelings of self-worth directly, as opposed to a downstream consequence of threat (defensiveness). This eliminates the concern that our investigation may teach us more about why threats have the ultimate effects they do, not why affirmations reduce threat. Nonetheless, to connect our research to examinations of how affirmations reduce defensiveness, our remaining experiments examined defensiveness directly.

Experiment 2

In Experiment 2, some participants self-affirmed before receiving false personality feedback in the form of 36 statements, 24 of which were negative. Participants were able to self-pace the administration of the feedback. As a measure of defensiveness, we measured how long participants exposed themselves to the negative feedback. Sedikides and Green (2000) showed that when processing this feedback with reference to the self (vs. a stranger), people devoted less time and processing resources to considering the negative behaviors. The authors inferred this indirectly, by showing that a memory advantage for the positive over negative feedback was eliminated when the feedback was displayed for limited time (2 s each) versus ample time (8 s each). Based on this research, we predicted that defensive participants would spend less time looking at the unfavorable feedback.

After seeing the feedback, participants made two judgments (modeled after those in the pilot study) that assessed whether they viewed the threat with perspective, and whether they trivialized it. We predicted that (a) self-affirmed participants would display less defensiveness by considering the negative feedback for longer than would control participants, (b) self-affirmed participants would have greater perspective on the feedback by considering it to be more narrow (but not trivialize it more), and (c) perspective (but not trivialization) would mediate the effect of affirmation on reduced defensiveness.

Method

Participants and design. Ninety-five Cornell University undergraduates were randomly assigned to an affirmation or control condition. We excluded 1 participant who indicated suspicion during debriefing that his personality feedback was actually false.

Procedure. Participants in the affirmation condition began by completing the values-based affirmation task used in Experiment 1. Control participants instead completed filler questionnaires, used in prior self-affirmation research (Critcher et al., 2010) that asked them to rate how appealing different jelly bean flavors and candle scents seemed. Next, participants were reminded that they had completed a number of personality scales when they signed up for the study (e.g., the Ten-Item Personality Inventory [TIPI]; Gosling, Rentfrow, & Swann, 2003). Participants were told they had completed the Durham Inventory for Behavioral Expectancies or "DIBE," which was able to provide feedback in a unique form. Supposedly, the inventory could predict how the participants—compared with their peers—were likely to behave, or how others were especially likely to behave toward the participants. At that point, the experimenter logged into a computer program using the participant's ID number, which appeared to link to participants' responses on the DIBE.

Participants were told that the behavioral expectancies would appear one at a time on the computer screen. To see the next one, they could press the space bar. In actuality, all participants saw the same 36 behaviors (from Sedikides & Green, 2000). Twenty-four represented negative feedback (e.g., "An employer would not rely on you to have an important task completed by the deadline."). The expected behaviors—24 negative, 12 positive—appeared in a random order. Based on the empirical precedent of Sedikides and Green (2000), we reasoned that defensiveness would be reflected by participants' devoting minimal looking time to negative feedback. We included the 12 positive behaviors as well to (a) make the feedback seem more realistic and (b) allow us to control for looking time.

After all feedback had been viewed, participants answered two final questions that had been modified from the items used in our pilot study (given the differences in what was being judged). One item assessed perspective: "If one received negative feedback on the DIBE, are these behaviors specific enough that other aspects of a person could overcome these limitations?" The other item assessed trivialization: "How important are the domains covered by the DIBE?" (reverse-scored). Each was responded to on a scale from 1 (*not at all*) to 9 (*completely*). We worried that completing the perspective item might itself lead people to adopt a broader perspective on the self. For this reason, we measured perspective *after* our unobtrusive measure of defensiveness.

Results and Discussion

For each participant, we calculated the average number of milliseconds spent looking at each of the 24 negative statements. To control for individual differences in reading or general looking time at feedback, we also calculated the average time spent looking at the non-threatening, positive statements. We submitted the negative looking times to an ANCOVA, with the average time looking at the positive statements as a covariate. A significant (positive) influence of the covariate suggested there were predictable individual

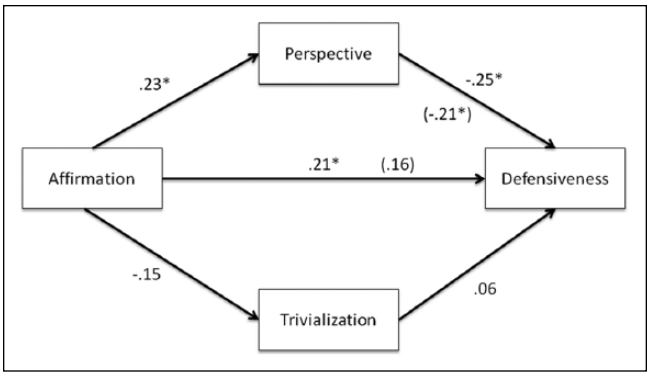

Figure 4. Perspective fully mediates the impact of self-affirmation on defensiveness reduction (Experiment 2).
Note. Self-affirmation does not operate similarly through trivialization. All numbers are standardized betas. The two betas in parentheses come from the regression model in which both pathways are estimated simultaneously.
*$p < .05$.

differences in how long participants spent looking at each statement, $F(1, 91) = 139.36$, $p < .001$, $\eta_p^2 = .60$. But, as expected, affirmed participants showed less defensiveness, spending more time looking at their negative feedback (M = 3.62s, SE = .08s) than did those in the control condition (M = 3.38s, SE = .08s), $F(1, 91) = 4.40$, $p = .04$, $\eta_p^2 = .05$, 95% CI = [0.06s, 0.55s].[7]

A 2 (affirmation) × 2 (item: perspective or trivialization) mixed-model ANOVA revealed the predicted interaction, $F(1, 93) = 6.14$, $p = .02$, $\eta_p^2 = .06$. Pairwise comparisons showed that affirmed participants had greater perspective on the threat (M = 6.74) than did those in the control condition (M = 5.94), $t(93) = 2.28$, $p = .03$, $d = .47$, 95% CI = [0.10, 1.51]. Affirmed participants were not more likely to trivialize the threat (M = 3.91) compared with those in the control condition (M = 4.33), $t(94) = -1.45$, $p > .15$, $d = -.30$, 95% CI = [−0.99, 0.16]. Similar to Simon et al.'s (1995) findings, those who showed greater perspective on the DIBE were not more likely to trivialize it. In fact, greater perspective was correlated with less trivialization, $r(93) = -.20$, $p = .05$. Those who displayed more perspective displayed less defensiveness, $pr(92) = -.22$, $p = .04$. In contrast, trivializing the feedback did not predict reduced defensiveness, $pr(92) = -.06$, *ns*.

To test whether perspective on the threat mediated the effect of affirmations on reduced defensiveness, we regressed the negative feedback looking time measure on the affirmation condition, perspective measure, and positive feedback looking time, simultaneously. Greater perspective continued to predict reduced defensiveness, β = .13, $t(90) = 2.00$, p = .05. The affirmation manipulation no longer did, β = .11, $t(90) = 1.64$, $p > .10$ (see Figure 4). We used Preacher and Hayes' (2008) bootstrapping method to test whether the indirect effect (the meditational pathway) was indeed significant. The 95% CI of the indirect effect through perspective did not include 0, [.0097, .2690]. Thus, perspective fully mediated the impact of self-affirmation on reduced defensiveness. In contrast, there was no evidence of a significant indirect effect through trivialization, [−.1035, .0556].

Experiments 3a and 3b

Our final pair of studies used an experimental approach to test whether perspective is responsible for the influence of self-affirmation on defensiveness. A standard values-based affirmation offers perspective on a threat, but it also does more. First, it prompts people to spend time writing about a valuable identity. Second, it selectively focuses people on a highly valued, positive domain at the expense of the threatened or less important domains. Third, it encourages people to engage with positive self-aspects by searching for meaning in those identities and analyzing why they contribute to one's worth as a person.

Is it specifically perspective, or is it instead one or more of these additional components of a values-based affirmation (perhaps only when combined with perspective) that are responsible for self-affirmations' effects? In our final experiments, we design a new manipulation that offers the same perspective as a standard self-affirmation but removes or substantially reduces the other three dimensions. We then test whether the influence of a standard self-affirmation is redundant with our new perspective exercise.

For this perspective exercise, participants drew in multiple identities—the threatened one, their most important identity (the sole focus of a typical self-affirmation manipulation), and their least important identity (the sole focus of a standard control exercise)—in a visual representation of their working self-concept. This task involves no written analysis, nor does it selectively focus people on their most important identity. But this task does offer perspective on the threat. It, quite literally, prompts participants to expand the working self-concept by recognizing additional identities in the self.

Experiment 3a not only has the potential to confirm that affirmational thinking and perspective are empirically distinguishable but also the study tests whether our perspective exercise has the intended effect—selectively disengaging affirmational processes without influencing perspective. Experiment 3b then tests whether the perspective exercise reduces defensiveness in a manner redundant with a standard self-affirmation. In other words, if a standard self-affirmation reduces defensiveness because it offers perspective on a threat—the element that the perspective exercise retains—then completing one or both of these exercises should have a similar defensiveness-reducing effect compared with completing neither. If a standard self-affirmation exercise reduces defensiveness instead because it prompts people to search for meaning in valued self-aspects and analyze why they contribute to a sense of self-worth (the affirmational thinking that the standard values-based affirmation engages), then the self-affirmation should have a unique (or at least superior) effect in reducing defensiveness.

Experiment 3a

Method

Participants. Participants were 103 undergraduates at the University of California, Berkeley, who took part in a 2 (exercise: perspective circle or self-affirmation) × 2 (measure: perspective or affirmation) fully within-subjects design.

Procedure. To first focus people on threat (and to mimic how Experiment 3b will open), participants recalled a time in which their performance on a test or assignment in their academic major did not live up to their own "personal academic standards." Next, participants ranked the personal importance of eight non-academic identities (e.g., athletic achievement). At this point, participants completed a self-affirmation (as used in Study 2) or our newly engineered perspective exercise (described below) in a counterbalanced order.

Perspective exercise. Participants were presented with an unshaded circle of 6.1 cm radius. One radius was already drawn. Participants were asked to imagine that the circle represented who they were as a person. Given that a person's identity is defined by any number of facets or aspects, we told them that different wedges of the circle corresponded to different aspects of their identity. Participants were then told that they would mark off three separate "wedges" on their self-circle to represent the "size" of three distinct aspects of their identities. We emphasized that the size of each wedge should correspond to its importance to the self and pointed out it was unlikely the three wedges would fill the entire circle.

Participants first considered their academic major. They were to move clockwise from the provided radius and draw a second radius to partition off the part of the self-concept that reflected the importance of their academic identity "that values success in [their] academic major." Next, participants continued clockwise around the perspective circle to partition off their most valued identity and their least valued identity as indicated by their earlier rankings. Participants labeled their wedges "major," "first," and "eighth" (Figure 5).

Did the exercises prompt perspective and/or affirmational thought? After completing both exercises, participants completed 14 items, once for each exercise. Seven items asked to what extent the exercise prompted perspective on the threatened identity (e.g., "made me think of one or more aspects of myself [beyond my academic self]"). Seven items assessed whether the exercise prompted people to affirm the self (e.g., "led me to search for a sense of meaning"). Instructions presented before these measures clarified that participants should report on "what the exercise prompted you to think about or do, regardless of whether the instructions explicitly instructed you to or you did so spontaneously." All responses were made on 7-point scales anchored at 1 (*no, not at all*) and 7 (*yes, quite clearly so*). The midpoint 4 was labeled "*somewhat*."

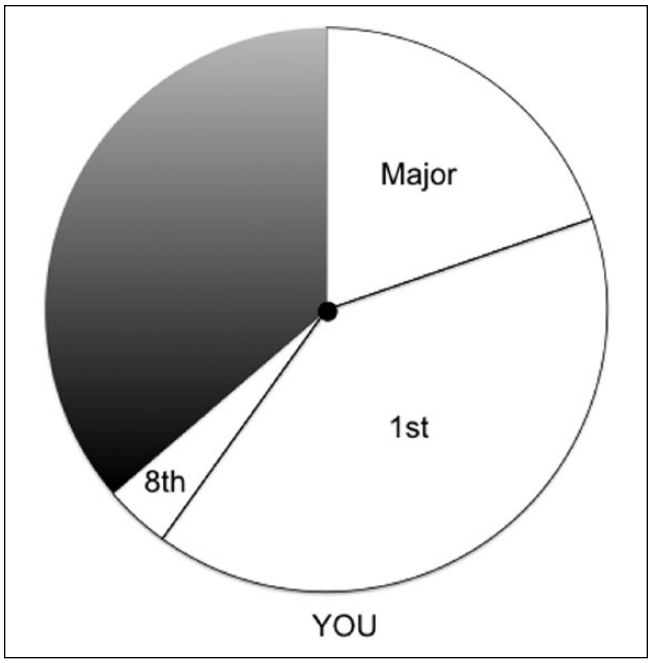

Figure 5. The perspective exercise completed with the average-sized identity wedges.
Note. Major = academic major; 1st = most valued identity; 8th = least valued identity. The dark space was left empty (Experiment 3).

Results

First, we wanted to know whether our measures of perspective and affirmational thought were indeed distinct. For both the 14 items measuring the experience of the writing exercise (the standard self-affirmation) as well as the 14 items measuring the experience completing the perspective circle, we performed principal components analyses with varimax rotation. In both cases, the expected (identical) two-factor solutions emerged. Items always loaded on the intended factor ($.58 < \lambda s < .87$) and less so on the other factor (see Appendix B). We averaged the items, as planned, to compute separate composites for perspective and affirmational processes for both the standard self-affirmation exercise as well as the perspective exercise ($.89 < \alpha s < .94$).

Second, we tested whether the self-affirmation exercise and the perspective circle engaged affirmation and perspective processes to different extents. We submitted the composites to a 2 (exercise: perspective circle or self-affirmation) × 2 (measure: perspective or affirmational thought) repeated-measures ANOVA, with only the second factor measured within-subjects. The predicted interaction significantly emerged, $F(1, 102) = 50.96$, $p < .001$, $\eta_p^2 = .06$. As expected, the standard self-affirmation exercise prompted as much perspective ($M = 4.91$, $SD = 1.33$) as did the perspective exercise ($M = 4.82$, $SD = 1.37$), $t < 1$; however, the self-affirmation prompted more affirmation ($M = 4.98$, $SD = 1.33$) than did the perspective exercise ($M = 3.82$, $SD = 1.33$), $t(102) = 7.96$, $p < .001$, $d = .78$, 95% CI = [0.87, 1.45]. Analyzed

differently, whereas the self-affirmation prompted similar reports of affirmation and perspective, $t < 1$, the perspective circle encouraged more self-reported perspective than affirmational thought, $t(102) = 7.88$, $p < .001$, $d = .78$, 95% CI = [0.75, 1.25].

Experiment 3b

Method

Participants and design. Participants were 174 Cornell University undergraduates who were randomly assigned to one of four conditions of a 2 (affirmation) × 2 (perspective exercise), between-participants design.

Procedure. As in Experiment 3a, participants first focused on a threat to their academic self by recalling a time in which their performance on a test or assignment in their academic major did not live up to their own "personal academic standards." At this point, participants completed one of two versions of the self-affirmation exercise. Those in the self-affirmation condition completed the same values affirmation task that was used in Experiments 2 and 3a. Those in the no-affirmation condition also ranked the eight domains, but wrote instead about their least valued domain and why it might be important in someone else's life.

Participants then received the perspective exercise and the defensiveness items in one of the two possible orders. In this way, we could vary whether the perspective exercise could influence defensiveness (i.e., prospectively), but still have everyone complete the measure to use it as a measure of trivialization. That is, we measured the size of each wedge to understand whether self-affirming changed the size or perceived importance of the threatened identity.

Defensiveness. We measured whether participants showed a defensive attribution style in explaining away academic failure. Because past research has found that defensiveness is more likely to be blocked in prospect than undone in retrospect (Critcher et al., 2010), we did not want to ask people to explain why they failed at the episode they recalled. Instead, we asked participants to imagine that on their next important assignment in their major, they did not do as well as they would have hoped. We then asked them to indicate the most likely explanation for such a prospective failure. Participants then expressed their agreement with three explanations, all of which were meant to detect a defensive attribution style ($\alpha = .59$): "The poor feedback didn't speak to my abilities, but to a set of unfortunate circumstances," "The failure was caused by circumstances associated with my external environment," and "My evaluator was unfair in his or her assessment of my work." The modesty of the reliability was largely attributable to the length of the scale: All items loaded on a single factor, $\lambda s > .67$. Responses were made on 9-point scales.

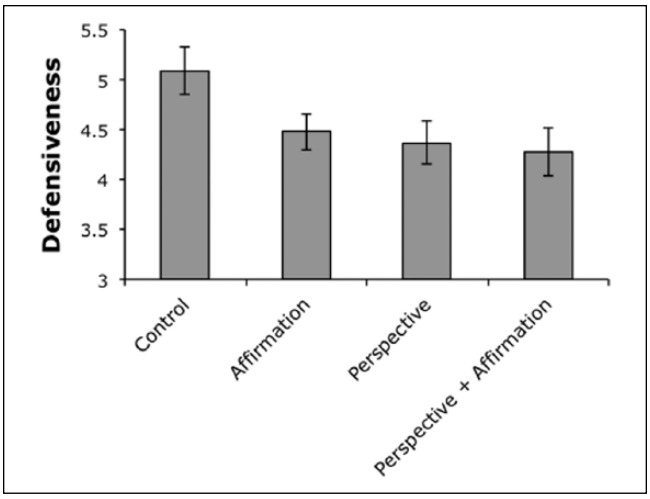

Figure 6. The perspective exercise, the self-affirmation, and the two in concert, all reduce defensiveness.
Note. Higher defensiveness scores reflect a more self-serving attributional style in explaining future academic failure. Means are the average of the three defensiveness items. (Experiment 3).

Results and Discussion

Defensiveness. We predicted that the self-affirmation alone (−1), the perspective exercise alone (−1), or both interventions in combination (−1) would reduce defensiveness similarly compared with completing neither (+3). This key contrast emerged, $t(168) = 2.75$, $p = .01$, $d = .42$, 95% CI = [0.20, 1.23], (Figure 6). The residual variance was non-significant, $F < 1$. Specific comparisons found that those who received neither manipulation showed greater defensiveness than those who self-affirmed only, $t(168) = 1.99$, $p = .05$, $d = .31$, 95% CI = [0.00, 1.23]; those who completed the perspective circle only, $t(168) = 2.28$, $p = .02$, $d = .35$ 95% CI = [0.10, 1.36]; and those who completed both exercises, $t(168) = 2.57$, $p = .01$, $d = .40$, 95% CI = [0.19, 1.44]. Completing either exercise, or the two in combination, all reduced defensiveness to the same extent, $ts < 1$. It is instructive not only that the perspective exercise was as effective as the self-affirmation in reducing defensiveness but also that the effects were not additive; this supports our contention that the process they operate through is redundant. Combined with the results of Experiment 3a (as well as Experiments 1 and 2), we can most parsimoniously conclude that it is self-affirmation's and the perspective exercise's similarity (perspective) instead of their substantial difference (affirmational thought) that explains their similar and redundant effects.

Trivialization. We measured the size of the drawn wedges to determine whether self-affirming changed the size or importance of these identities. By our account, affirmations only fill in additional important identities in the self-concept; they do not lead people to trivialize the importance

(which size was to indicate) of threatened identities. Using a protractor, a coder measured the number of degrees of each of the three wedges for each participant. We submitted a number of measures to 2 (affirmation) × 2 (perspective exercise: pre- or post-defensiveness measure) ANOVAs: the size of the academic major ($M = 71.07°$, $SD = 45.67°$), the most valued identity ($M = 144.61°$, $SD = 57.36°$), and the least valued identity ($M = 13.74°$, $SD = 13.93°$) wedges. Their sizes did not vary by condition, Fs < 1.62, ps > .20, η_p^2s < .01. It appears that completing a standard values-based affirmation or a perspective exercise may reduce defensiveness by "filling in" other aspects of the working self-concept without changing the importance of these identities.

General Discussion

An abundance of research has shown that self-affirmation has a wide range of effects on emotion, cognition, and behavior (Sherman & Cohen, 2006; Steele, 1988), often reducing or eliminating defensive reactions to psychological threat. What is less well understood is why self-affirmation has this wide range of effects. We proposed that threats constrict the working self-concept to focus on threatened self-aspects, whereas self-affirmations expand the working self-concept, thus offering broader perspective on the threat. As a result, the threatened identity looms less evaluatively large, thereby failing to evoke defensive responses. In this way, self-affirmation does not cause one to trivialize the threat or threatened identity; it simply reminds people and their working self-concepts that additional important aspects of the self exist.

Experiment 1 showed that self-affirmations expand a self that has become constricted under threat. Without affirmation, a threatened identity dominated feelings of self-worth. However, a self-affirmation broadened participants' perspective, thereby weakening the evaluative impact of the threat and restoring a positive sense of integrity based on one's broader views of self. Experiment 2 found that the self-affirmed reported greater perspective on, but not trivialization of, the threat. This enhanced perspective mediated the effect of self-affirmation on reduced defensiveness. Experiments 3a and 3b found that manipulating the size of the working self-concept directly by means of a "perspective exercise"—completing a visual representation of one's working self-concept to include multiple identities—prompted perspective (but not affirmational thinking) and, in turn, reduced defensiveness. That is, an exercise that prompted perspective on a threatened identity, but did not encourage writing about a valued identity, focusing solely on a valued identity, or analyzing why a valued identity contributes meaning and worth to one's life, reduced defensiveness in a manner that was similar to (and redundant with) a standard self-affirmation exercise.

Relation to Other Work

The influence of threat on global feelings of self-worth has not been discussed in terms of a constricted self-concept that follows threat, but has been in terms of overgeneralization of the threat itself (Crocker & Park, 2004). For example, Tangney, Wagner, Hill-Barlow, Marschall, and Gramzow (1996) discussed how some negative behavior or failures are taken to reflect a more global self-defect (see also Brown & Dutton, 1995). Our approach, instead, highlights the importance of the working self-concept and suggests that threats do not necessarily expand to contaminate one's view of the entire self. Instead, the self "expands" after self-affirmation to make a threat seem narrow.

We think these are not merely two ways of describing the same effect. If threat were indeed overgeneralized, it would not be clear why providing a broad perspective on the self via self-affirmation would be effective in restoring feelings of self-worth. Instead, self-affirmation would only call to mind more identities into which the threat had bled. The language of overgeneralization implicitly holds the size of the self fixed while considering how threats can grow larger or smaller. We instead note that under threat, the size of the working self may constrict, an effect that can be undone by self-affirmation. Thus, although both low self-esteem individuals (Kernis, Brockner, & Frankel, 1989) and depressives (Beck, 1967; Carver, Ganellen, & Behar-Mitrani, 1985) have been said to overgeneralize from failure experiences, future research may wish to test whether such individual differences reflect overgeneralization or instead a difficulty regaining perspective on the broader self.

More generally, our approach suggests that the best way to deal with threat is not necessarily to engage in extensive consideration of or rumination about a threat. Such obsessions only serve to maintain a narrowly focused working self-concept. This lesson resembles a principle that has emerged from work on rumination and aggression. Ruminating about a threat, compared with distracting oneself, may actually heighten threat-related responses (Bushman, Bonacci, Pedersen, Vasquez, & Miller, 2005). And in fact, Koole et al. (1999) showed that self-affirmations reduced rumination about a failure. By our account, the defocusing responsible for the ending of rumination is itself a demonstration of perspective. That said, the affirmation-as-perspective approach also makes it clear why affirmations do not work as mere distractions (Steele, 1988). For when one's attention returns to the self, the working self-concept will still be constricted and threatened unless something (like a self-affirmation) has expanded the self and thus offered perspective on the threat.

One reading of our approach is that affirmations shift situational contingencies of self-worth. Crocker and Wolfe (2001) describe that self-evaluations rise and fall with the successes and failures of some identities (contingent ones)

but not others. Typically, contingencies of self-worth are thought of as stable associations that are dominant in guiding self-evaluations. In our story, a broad or narrow perspective on the self may change one's situational contingencies of self-worth. Immediately after threat, a threatened identity will loom disproportionately large, and self-worth will be especially contingent on it (see pilot study and Experiment 1). Self-affirmations increase the accessibility of more of the self (Croizet, Desert, Dutrevis, & Leyens, 2001), leading self-worth to be no longer as situationally contingent on any single self-aspect.

Ability to Account for Previous Findings

The current approach also sheds light on why certain affirmations are not effective. By our account, an affirmation must expand the working self-concept to bolster the self, making multiple identities active in a person's working self-concept. Thus, it should not be a surprise that affirming someone in the same domain in which they are threatened is typically ineffective, and can even produce backfiring defensiveness-enhancing effects (Blanton, Cooper, Skurnik, & Aronson, 1997; Sivanathan, Molden, Galinsky, & Ku, 2008). Such affirmations only reinforce a constricted working self-concept focused squarely on the threatened domain. Given the parallels in our reasoning to the self-complexity literature, it may be the case that affirming identities that are not only different but also especially distinct or differentiated from the threatened identity may be most effective (Linville, 1985, 1987). This nuanced prediction awaits future research.

Although one must focus on more than the non-threatened identity to achieve perspective on it, does it matter whether one focuses on relatively positive, negative, or neutral identities? Expanding the working self-concept into additional threatened domains does not inject the working self-concept with (to use the language of our perspective measure from Experiment 2) compensating sources of self-evaluation, explaining why such manipulations heighten defensiveness (Cohen et al., 2007). Such self-expansion merely compounds the sources of threat. But does the affirmation as perspective account demand that the affirmed identities be positive? Given people tend to deemphasize self-aspects they feel less positively about (Pelham & Swann, 1989), positive identities should be more effective in providing the most perspective-offering self-concept expansion. And as Experiment 3b showed, focusing on an important, sizable identity reduced defensiveness more than focusing on an unimportant one. But Experiment 3b also showed that calling to mind both positive and more neutral or negative self-aspects (the perspective exercise) was as effective in reducing defensiveness as merely focusing on an important, positive self-aspect (the self-affirmation exercise). This suggests that it may be expansion, not the average positivity of self-relevant information that is brought to mind, that is key. In Experiment 3b, we found that participants drew their most important identity as more than 10 times larger than their least important identity. Thus, future research could test whether having participants focus on 10 relatively unimportant identities reduces threat-inspired defensiveness as effectively as affirming a single important identity. If so, it would show the (relatively inefficient) way that less-positive identities could offer defensiveness-reducing perspective.

Not only does our account help identify boundaries on affirmations' effectiveness, it also helps explain why certain affirmations are surprisingly effective. Our account argues (and Experiment 3b showed directly) that expanding the working self-concept to include compensating, perspective-offering identities is sufficient to reduce defensiveness. There is no reason to expect that such expansion need be done consciously. And indeed, Sherman, Cohen, et al. (2009) found that implicitly priming an identity by way of a sentence-unscrambling task served to alleviate a threat response.

The present approach also helps explain why certain manipulations that do not look like typical self-affirmations appear to reduce defensiveness in the same way. For example, completing a self-esteem scale—thereby prompting a focus on one's broad resources—serves as a self-affirmation for those with high self-esteem (Steele, Spencer, & Lynch, 1993). Furthermore, our approach makes more understandable a mystery noted by Steele et al. (1993)—that high self-esteem people did not spontaneously call on their broader resources in dealing with a threat. By recognizing that threats promote narrow perspective and that disengagement is difficult (Koole et al., 1999; Koster et al., 2004; Rydell et al., 2009), it is understandable that it may take an external intervention (e.g., a self-affirmation, a perspective exercise) to promote defocusing and broader perspective.

The affirmation as perspective account also explains why completing an Allport-Vernon-Lindzey Values scale related to an important identity serves as a self-affirmation (Allport, Vernon, & Lindzey, 1960; Tesser & Cornell, 1991), why completing the same scale as though one did not care about the value still serves as an affirmation, but why completing a scale about an unimportant value does not (Simon et al., 1995). In completing these scales, one endorses statements related to certain values or identities. According to the affirmation as perspective model, defensiveness reduction does not come from value expression itself, but from making accessible a sizable, but otherwise dormant aspect of one's self-concept. Of course, if the value scale does not relate to a valued aspect of one's identity, then no working self-concept expansion is likely, explaining why such value expression did not serve as an effective affirmation.

Note that the logic advanced here to understand the impact of self-affirmation can be understood in the context of how other sources of threats are ultimately downplayed. Miron, Branscombe, and Biernat (2010) showed that Americans are motivated to "raise the standards" for what constitutes racism to downplay the threatening possibility that the United

States' troubling experience with race implies that it is a racist nation. But after listing three positive things that Americans had done over the course of history and explaining why those three things were particularly characteristic of America, this tendency was reduced. Although the manipulation differs from self-affirmation manipulations in that it does not directly bolster the self, it does provide perspective on one's view of America. It defocuses one from America's racial injustices, allowing a person to acknowledge a pocket of injustice without having to abandon one's overall positive accounting of one's country.

Conclusion

Sherman and Cohen (2006) argued that it is unlikely that any single mediator will account for all effects of self-affirmation on defensiveness. We agree with this view, for mediators can tell as much (if not more) about how threat leads to defensive responding as about how affirmations reduce threat. Furthermore, given the variety of ways in which self-affirmation has been operationalized (McQueen & Klein, 2006), it may be that different affirmation techniques have their effects for somewhat different reasons. Crocker et al. (2008) argued that affirmations may "remind people what they care about beyond themselves," thereby "transcending the self" in a way that induces "love," which may "inspire people to improve" (p. 741). We instead have argued that affirmations may remind people that they care about more than just the threatened identity and that by transcending these narrow concerns they may adopt proper perspective on the self. Both accounts emphasize that affirmations serve to deemphasize the implications of a threat by placing it in a broader context. Future affirmation research should seek to understand how and under what circumstances affirmations can best achieve perspective.

Appendix A

The 14 feelings of self-worth items (Experiment 1)—responded to on scales from 1 (*not at all*) to 9 (*extremely*)—loaded on two distinct factors that were identified by an exploratory principal components analysis with varimax rotation. The items, and their respective factor loadings, are listed below:

Positive Feelings of Self-Worth

I currently feel proud. (.80)
I currently feel confident. (.77)
Overall, I feel positively toward myself right now. (.75)
I feel like a successful individual. (.74)
I currently feel pleased with self. (.73)
I feel good about myself right now. (.69)
I feel very much like a person of worth. (66)
I do not feel very confident in myself right now. (−.63)

Negative Feelings of Self-Worth

I currently feel uneasy. (.83)
I currently feel humiliated. (.82)
I current feel ashamed. (.82)
I currently feel bothered. (.79)
I feel inferior at this moment. (.73)
I am frustrated or rattled. (.63)

Appendix B

The following items assessed the extent to which people adopted perspective (a recognition that the self is made up of more than just the threatened identity) or engaged in affirmational thinking while completing each exercise in Experiment 3a. Following each item are two-factor loadings, reflecting the degree to which the item loaded on the intended factor when completed to describe one's experience with the perspective exercise or self-affirmation exercise, respectively.

Perspective

Made me think of one or more aspects of myself (beyond my academic self) (.87, .80)
Prompted me to identify additional (non-academic) aspects of my identity (.86, .68)
Thought of aspects of my identity that extend beyond academics (.85, .82)
Focused on one of more non-academic aspects of my identity (.82, .79)
Led me to identify one or more non-academic identities (.80, .85)
Offered perspective on my "academic self/identity" by reminding me of non-academic parts of who I am (.78, .67)
Led me to appreciate I have multiple parts of who I am (.71, .76)

Affirmational Thinking

Searched for a sense of meeting (.85, .64)
Emphasized why something has made my life affirming and whole (.81, .64)
Elaborated on why something has contributed to my worth as a person (.80, .84)
Constructed a narrative that describes what has made my life meaningful (.79, .83)
Sought to affirm my worth as a person (.76, .70)
Mulled over what makes me feel positive (.68, .70)
Dwelled on why something has been particularly meaningful to me (.59, .78)

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

Notes

1. In this way, dispositional self-esteem and self-assessment in the threatened domain are not direct measures of the working self-concept, but their differential ability to account for one's situational sense of self-worth provides evidence that one's momentary sense of self is broad or narrow, respectively.
2. Because affirming the threatened domain would not expand the self and offer perspective on the threat, and given previous research has found that affirming the self in the threatened domain tends not to be effective in reducing defensiveness (Blanton, Cooper, Skurnik, & Aronson, 1997), participants were not permitted to affirm their academic or intellectual self in this study, or the analogous threatened domains in the subsequent studies.
3. Readers may note that not only did our control condition participants not complete an affirmation but they also did not see the list of additional identities. According to our account (and as Experiment 3b more directly tests), it is the expansion of the working self-concept that comes from activating additional identities, not the affirmation exercise itself, that is crucial. Thus, letting participants view a list of additional identities might be akin to an affirmation condition. In fact, Sherman, Cohen, et al. (2009) found this was the case: Priming identities served as a self-affirmation. Sherman et al.'s findings suggest that many self-affirmation effects may in large part be driven by the control condition in which people are exposed to multiple identities, but then are asked to focus on a particularly unimportant identity; this may be focusing people on a threat, given the self tends to devalue those aspects about which it feels less confident, competent, or secure. Because our goal is to understand why self-affirmations are effective, not how the standard control conditions may also contribute to affirmation effects, Experiments 1 and 2 do not use that typical control. At the same time, to assure readers that our effects are not dependent on this methodological decision, we used a more typical control condition in Experiment 3b. By observing consistent effects in our studies, we can be more assured that support for our approach does not depend on one methodological approach.
4. Lion.
5. Note that these predictors of situational feelings of self-worth—dispositional self-esteem and self-assessment in the threatened domain—differ not only in their content but also in two additional ways. First, they were administered at different points in time. We assessed dispositional self-esteem in advance because (a) we wanted to get a baseline measure of self-esteem before one's working self-concept had been constricted by threat, and (b) had we measured self-esteem in the lab before the threat, it could have acted as a self-affirmation manipulation (Steele, Spencer, & Lynch, 1993). Second, the measures differed in number of items: Dispositional self-esteem was assessed with a 10-item scale; self-assessment in the threatened domain, with a single item. Each difference may have improved or degraded the measure's ability to capture a true relationship between the measured predictor and the measured outcome (feelings of self-worth) but would not account for why their relative predictive power differed by affirmation condition.
6. Even though there was no main effect of affirmation condition on negative feelings of self-worth, the same evidence of perspective emerged on negative feelings of self-worth as well. Notably, the test for the difference between the betas again showed that affirmed participants had broader perspective, $t(64) = 2.38$, $p = .02$. Providing support for our affirmation as perspective account, affirmed participants' negative feelings of self-worth were more tethered to their broad self-assessment versus narrow self-assessment ($\beta = -.37$ vs. $\beta = -.09$); the non-affirmed, to their narrow self-assessment versus broad self-assessment ($\beta = -.48$ vs. $\beta = -.05$).
7. Using the positive statement looking time as a covariate presupposes that it is a measure of reading speed or looking time that is not influenced by the affirmation manipulation. That is, if the self-affirmation merely increased interest in all feedback (instead of encouraging people to look more at negative feedback in particular), then the effects on negative looking time may have emerged merely because the covariate was imperfect in controlling for this source of shared variance. Countering this possibility, the affirmation manipulation did not affect reading time on the positive statements, $t < 1$. And even when no effort was made to control for individual differences in looking time, there still emerged a (marginally) significant effect of affirmation condition on the negative statements, $t(92) = 1.83$, $p = .07$.

Supplemental Material

The online supplemental material is available at http://pspb.sagepub.com/supplemental.

References

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

Beck, A. T. (1967). *Depression: Clinical, experimental, and theoretical aspects*. New York, NY: Harper & Row.

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*, *23*, 684-696.

Brown, J. D., & Dutton, K. A. (1995). The thrill of victory, the complexity of defeat: Self-esteem and people's emotional reactions to success and failure. *Journal of Personality and Social Psychology*, *68*, 712-722.

Bushman, B. J., Bonacci, A. M., Pedersen, W. C., Vasquez, E. A., & Miller, N. (2005). Chewing on it can chew you up: Effects of rumination on triggered displaced aggression. *Journal of Personality and Social Psychology*, *88*, 969-983.

Carver, C. S., Ganellen, R. J., & Behar-Mitrani, V. (1985). Depression and cognitive style: Comparisons between measures. *Journal of Personality and Social Psychology*, *49*, 722-728.

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

Cohen, G. L., Sherman, D. K., Bastardi, A., Hsu, L., McGoey, M., & 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.

Critcher, C. R., Dunning, D., & Armor, D. A. (2010). When self-affirmations reduce defensiveness: Timing is key. *Personality and Social Psychology Bulletin*, *36*, 947-959.

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.

Crocker, J., & Wolfe, C. T. (2001). Contingencies of self-worth. *Psychological Review*, *108*, 593-623.

Croizet, J., Desert, M., Dutrevis, M., & Leyens, J. (2001). Stereotype threat, social class, gender, and academic under-achievement: When our reputation catches up to us and takes over. *Social Psychology of Education*, *4*, 295-310.

Dixon, T. M., & Baumeister, R. F. (1991). Escaping the self: Moderating effects of self-complexity. *Personality and Social Psychology Bulletin*, *17*, 363-368.

Dunning, D., Leuenberger, A., & Sherman, D. A. (1995). A new look at motivated inference: Are self-serving theories of success a product of motivational forces? *Journal of Personality and Social Psychology*, *59*, 58-68.

Epton, T., & Harris, P. R. (2008). Self-affirmation promotes health behavior change. *Health Psychology*, *27*, 746-752.

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.

Gosling, S. D., Rentfrow, P. J., & Swann, W. B. Jr. (2003). A very brief measure of the Big Five personality domains. *Journal of Research in Personality*, *37*, 504-528.

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.

Howell, J. L., & Shepperd, J. A. (2012). Reducing information avoidance through affirmation. *Psychological Science*, *23*, 141-145.

Karpinski, A., & Steinman, R. B. (2006). The single category implicit association test as a measure of implicit social cognition. *Journal of Personality and Social Psychology*, *91*, 16-32.

Kernis, M. H., Brockner, J., & Frankel, B. S. (1989). Self-esteem and reactions to failure: The mediating role of overgeneralization. *Journal of Personality and Social Psychology*, *57*, 707-714.

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.

Koster, E. H. W., Crombez, G., Verschuere, B., & De Houwer, J. (2004). Selective attention to threat in the dot probe paradigm: Differentiating vigilance and difficulty to disengage. *Behaviour Research and Therapy*, *42*, 1183-1192.

Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin*, *108*, 480-498.

Linville, P. W. (1985). Self-complexity and affective extremity: Don't put all of your eggs in one cognitive basket. *Social Cognition*, *3*, 94-120.

Linville, P. W. (1987). Self-complexity as a cognitive buffer against stress-related illness and depression. *Journal of Personality and Social Psychology*, *52*, 663-676.

Markus, H., & Wurf, E. (1987). The dynamic self-concept: A social psychological perspective. *Annual Review of Psychology*, *38*, 299-337.

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.

Mathews, A., Mackintosh, B., & Fulcher, E. P. (1997). Cognitive biases in anxiety and attention to threat. *Trends in Cognitive Sciences*, *9*, 340-345.

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

Mednick, S. A. (1962). The associative basis of the creative process. *Psychological Review*, *69*, 220-232.

Miron, A. M., Branscombe, N. R., & Biernat, M. (2010). Motivated shifting of justice standards. *Personality and Social Psychology Bulletin*, *36*, 768-779.

Mogg, K., & Bradley, B. P. (1999). Orienting of attention to threatening facial expressions presented under conditions of restricted awareness. *Cognition & Emotion*, *13*, 713-740.

Niedenthal, P. M., Setterlund, M., & Wherry, M. B. (1992). Possible self-complexity and affective reactions to goal-relevant evaluation. *Journal of Personality and Social Psychology*, *63*, 5-16.

Pelham, B. W., & Swann, W. B. Jr. (1989). From self-conceptions to self-worth: On the sources and structure of global self-esteem. *Journal of Personality and Social Psychology*, *57*, 672-680.

Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior Research Methods*, *40*, 879-891.

Rosenberg, M. (1965). *Society and the adolescent self-image*. Princeton, NJ: Princeton University Press.

Rydell, R. J., McConnell, A. R., & Beilock, S. L. (2009). Multiple social identities and stereotype threat: Imbalance, accessibility, and working memory. *Journal of Personality and Social Psychology*, *96*, 949-966.

Sedikides, C., & Green, J. D. (2000). On the self-protective nature of inconsistency/negativity management: Using the person memory paradigm to examine self-reference memory. *Journal of Personality and Social Psychology*, *79*, 906-922.

Sherman, D. K., Bunyan, D. P., Creswell, J. D., & Jaremka, L. M. (2009). Psychological vulnerability and stress: The effects of self-affirmation on sympathetic nervous system responses to naturalistic stressors. *Health Psychology*, *28*, 554-562.

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., Cohen, G. L., Nelson, L. D., Nussbaum, A. D., Bunyan, D. P., & Garcia, J. (2009). Affirmed yet unaware: Exploring the role of awareness in the process of self-affirmation. *Journal of Personality and Social Psychology*, *97*, 745-764.

Sherman, D. K., & Hartson, K. A. (2011). Reconciling self-defense with self-criticism: Self-affirmation theory. In M. D. Alicke &

C. Sedikides (Eds.), *Handbook of self-enhancement and self-protection* (pp. 128-151). New York, NY: Guilford Press.

Simon, L., Greenberg, J., & Brehm, J. (1995). Trivialization: The forgotten mode of dissonance reduction. *Journal of Personality and Social Psychology*, *68*, 247-260.

Sivanathan, N., Molden, D. C., Galinsky, A. D., & Ku, G. (2008). The promise and peril of de-escalating commitment through self-affirmation. *Organizational Behavior and Human Decision Processes*, *107*, 1-14.

Spencer, S. J., Zanna, M. P., & Fong, G. T. (2005). Establishing a causal chain: Why experiments are often more effective in examining psychological process than mediational analyses. *Journal of Personality and Social Psychology*, *89*, 845-851.

Steele, C. M. (1975). Name-calling and compliance. *Journal of Personality and Social Psychology*, *31*, 361-370.

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

Steele, C. M., Spencer, S. J., & Lynch, M. (1993). Self-image resilience and dissonance: The role of affirmational resources. *Journal of Personality and Social Psychology*, *64*, 885-896.

Stone, J., Whitehead, J., Schmader, T., & Focella, E. (2011). Thanks for asking: Self-affirming questions reduce backlash when stigmatized targets confront prejudice. *Journal of Experimental Social Psychology*, *47*, 589-598.

Tangney, J. P., Wagner, P. E., Hill-Barlow, D., Marschall, D. E., & Gramzow, R. (1996). Relation of shame and guilt to constructive versus destructive responses to anger across the lifespan. *Journal of Personality and Social Psychology*, *70*, 797-809.

Taylor, V. J., & Walton, G. M. (2011). Stereotype threat undermines academic learning. *Personality and Social Psychology Bulletin*, *37*, 1055-1067.

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

Wakslak, C. J., & Trope, Y. (2009). Cognitive consequences of affirming the self: The relationship between self-affirmation and object construal. *Journal of Experimental Social Psychology*, *45*, 927-932.

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology*, *54*, 1063-1070.