Self-Affirmation: Understanding the Effects

David K. Sherman
University of California, Santa Barbara

Geoffrey L. Cohen
Stanford University

ABSTRACT

Self-affirmation theory proposes that people possess a flexible self-system, such that they can respond to threats in one domain of life by affirming self-worth in other domains. In social psychology research, this has been examined in studies where people affirm important values in the context of self-threatening events or information. This paper reviews the literature demonstrating the effects of values affirmations and proposes a theoretical account to understand how self-affirmations reduce defensiveness in response to threats to individuals’ health, attenuate physiological stress responses to laboratory and naturalistic stressors, and improve academic performance among individuals experiencing identity threat. The proposed model has three components: Self-affirmations boost self-resources, broaden the perspective with which people view information and events in their lives, and lead to an uncoupling of the self and the threat, reducing the threat’s impact on the self. This model elucidates the experience and impact of affirming values in the context of threats and the processes through which self-affirmations trigger large and long-lasting effects.

SELF-AFFIRMATION: UNDERSTANDING THE EFFECTS

Over the past thirty years, psychological researchers have induced people to affirm their values in the context of self-threatening events and information. When affirmed, smokers are more open to anti-smoking information (Harris, Mayle, Mabbott, & Napper, 2007; Crocker, 1996); responsibility for their teams (Crocker & Kim, 2005), and minority students’ sense of belonging in school and showed gains over a period of time (Cohen, Garcia, Purdie-Vaughns, Garcia, & Holstein, 2006). The theory has evolved from an exploration of social identity phenomena (Steele, 1988; see also Crocker & Major, 2009) and informs intervention in a wide variety of contexts (Ehret & Sherman, 2014; Harms & Cohen, 2006 for reviews). The purpose of the present paper is (1) to review major new discoveries about how self-affirmations affect defensive reactions to group threat, and (3) present a general account of how affirming important values is available to an individual to combat stress, improve resources, a focal threat can be reduced so that people can see the events from a higher perspective, and to see them from a high enough level to experience the threat as personal. This theoretical account does not offer a general theory, but it offers a general picture of the self-affirmation process, so people affirm the self. As a result, people can re-appraise meanings across a disparate range of social and psychological experiences.

SELF-AFFIRMATION THEORY

Self-affirmation theory begins with the idea that the self is a flexible system and that one’s self can be adaptively adequate, and not threatened (Steele, 1988; see also Cohen & Sherman, 2006). Stressful or threatening situations can prompt rationalizations or the attribution of blame (Cohen & Sherman, 1991). These perceptions of the self are then transformed into threats to self-esteem, which then incites the victimization of others as described by another person who becomes the target of a self-affirming framework (Cohen & Sherman, 2014).
& Napper, 2007; Crocker, Niyi, & Mischkowski, 2008), athletes take more responsibility for their teams' defeats and less credit for their successes (Sherman & Kim, 2005), and minority students experiencing stereotype threat feel greater belonging in school and show improved academic performance over a substantial period of time (Cohen, Garcia, Purdie-Vaughns, Apfel, & Brzustoski, 2009; Cock, Purdie-Vaughns, Garcia, & Cohen, 2012; Sherman et al., 2013). Self-affirmation theory has evolved from an alternative explanation for cognitive dissonance phenomena (Steele, 1988; see also J. Aronson, Cohen, & Nail, 2009) to a theory that informs intervention in a wide range of settings (see Cohen & Sherman, 2014; Ehret & Sherman, 2014; Harris & Epton, 2009; Garcia & Cohen, 2012; Sherman & Cohen, 2006 for reviews). The question of precisely what it is that causes such effects is still — in many psychologists' eyes — unresolved and thus motivates continued research attention.

The purpose of the present review is to (1) summarize self-affirmation theory, (2) review major new discoveries in affirmation research with an emphasis on how affirmations affect defensiveness, stress, and academic performance under identity threat, and (3) present a general theoretical account of how these effects occur. We propose that affirming important values enhances the psychological resources available to an individual to confront a threat. With this enhanced perception of self-resources, a focal threat can be viewed from a broader perspective, as people tend to see the events from a higher level of construal. This broader perspective enables them to experience the threat such that its effect on self-evaluation weakens. This theoretical account does not spell out specific mediating mechanisms but instead offers a general picture of the psychological shift or mindset that emerges when people affirm the self. As a consequence, the account we offer applies well to findings across a disparate range of research studies, and to people across a wide range of social experiences.

SELF-AFFIRMATION THEORY

Self-affirmation theory begins with the premise that people are motivated to maintain the integrity of the self, a global perception of oneself as morally and adaptively adequate, and not their perceived worth in specific domains (Steele, 1988; see also Cohen & Sherman, 2014; McQueen & Klein, 2006; Sherman & Cohen, 2006). Stressful or threatening events — those that call into question perceptions of global adequacy — focus attention and motivate physiological and psychological resources to combat the perceived threat (Sapolsky, 2004; S. E. Taylor, 1991). These perceptions of environmental challenges to one's self-integrity can prompt rationalizations or other defensive responses aimed to lessen the threat (Cohen & Sherman, 2014). For example, being reminded of one's hurtful actions toward others could be threatening, but the threat could be reduced if one rationalizes the victims as deserving it (E. Aronson, 1999). Identity threats occur when an individual's self-view is challenged, and this includes stereotype threat, which arises in contexts when there is a possibility that a valued social identity
(e.g., ethnicity, gender, or nationality) could be devalued (Purdie-Vaughns, Steele, Davies, Dithmann, Randall Crosby, 2008; Steele, Spencer, & Arenson, 2002). Because they can be chronic and tightly linked to important life outcomes such as academic performance, identity threats have been a particular focus of self-affirmation researchers (Cohen, Garcia, Apfel, & Master, 2006; Cohen et al., 2009; Sherman et al., 2013).

What is central to affirmation theory is that the self-system is flexible, and people have many responses in their “psychological immune system” that they can draw on (Gilbert, Pinel, Wilson, Blumberg, & Wheatley, 1998). One such response is affirming the self in a domain altogether different from the instigating threat. When people affirm their overall self-integrity—that is, their view of the self as being capable and adaptive—they will have less need to defensively neutralize the threat through direct means, such as rationalizing it (Sherman & Cohen, 2006; Steele, 1988). Reminded of who they are and what is important to them, such self-affirmations can reduce stress by putting threats in the context of an overall narrative of self-integrity.

A self-affirmation is an act that demonstrates a person’s adequacy (Cohen & Sherman, 2008), and there are many experimental operationalizations of self-affirmation (McQueen & Klein, 2006). These include positive feedback on a personally important skill (e.g., Cohen, Arenson, & Steele, 2000, Study 2), purchasing of status goods (Sivanathan & Pettit, 2010), and updating one’s Facebook page (Toma & Hancock, 2013). This review focuses on values affirmations, as the focal outcomes reviewed have all been assessed most reliably as a function of values affirmations: moreover, other manipulations may operate through different processes (e.g., positive feedback may introduce mood effects that have generally not been observed to result from values affirmations). In addition, this review speaks to how values affirmations affect people under threat, as the process for those not experiencing threat is likely to be different (Briñol, Petty, Gallardo, & DeMarree, 2007).

In a standard affirmation induction, people write about central values or complete questionnaires that evoke their central values such as their relationships with friends and family. What is key about these manipulations is that they are self-generated and enable people to express what is important to them and why in their own words (Rokeach, 1973). When timed to threatening circumstances, such values affirmations can serve as turning points for an individual’s narrative, supporting adaptive coping under threat (Cohen & Sherman, 2014).

SELF-AFFIRMATION EFFECTS

Defensive Responses

As self-affirmation theory originated in the context of dissonance theory, early affirmation studies showed how people are less likely to rationalize their decisions if given the opportunity to affirm a key value even when unrelated to the decision (Steele & Liu, 1983; see key papers by Blanton, Cooper, Skurnik, & Arenson, 1997; Stone & Cooper, 2001). To thoughts should “reduce d & (p. 290; Steele, 1988). This l ening health information. A individual’s self-image both individual has acted in a mat whether bolstering the self i activities can reduce defen ifi studies is that people are de threats self-integrity. Peo by engaging in motivated i desired conclusion that the can be attenuated, then det (e.g., Reed & Aspinwall, 19 9 affirmed people in one stud ease (Howell & Spreer, 2 to taking a diabetes screeni There have now been m tion in health information st Epton, 2009; 2010; McQue see also Zhao, Peterson, K published in Health Psych for intentions and behavior, 95% confidence interval [0.12–1.16]) found that “Across 34 tests tions (N = 5,564), 46% dicated small but reliably acceptance, d = .17 (CI = .14 aor, d = .32 (CI = .19 to .44, 2015, p. 187).” Thus, there a self-affirmations in facilita health information, typically written or video form. Self- fical settings, with beneficial interventions with patients (see Cohen & Sherman, 2014).

In terms of mechanisms, processing, several important factors help people calibrate to and process, factors, such that people with a low risk are not unduly alarmed to greater message scrutiny (see Brick, Cohen, & Sherman, 2014).
Self-Affirmation: Understanding the Effects

Stone & Cooper, 2001). That initial work led to the hypothesis that self-affirming thoughts should “reduce defense mechanisms such as denial and rationalization” (p. 290; Steele, 1988). This hypothesis is highly relevant to how people process threatening health information. As health risk information has the potential to threaten an individual’s self-image both by linking the self with disease and by suggesting that an individual has acted in a maladaptive or harmful manner, many studies have examined whether bolstering the self in one domain by having people engage in value-affirming activities can reduce defensive, self-serving health assessments. The logic of these studies is that people are defensive in response to health risk information because it threatens self-integrity. People reduce the potential threat of the information, in part, by engaging in motivated inferences about the health information that leads to the desired conclusion that they are not at risk (Kunda, 1990). However, if self-threat can be attenuated, then defensiveness should be reduced and openness facilitated (e.g., Reed & Aspinwall, 1998; Sherman, Nelson, & Steele, 2000). For example, self-affirmed people in one study were more open to risk feedback for an untreatable disease (Howell & Shepperd, 2012). Affirmation also led at-risk people to be more open to taking a diabetes screening test (van Koningsbruggen & Das, 2009).

There have been many other demonstrations of the beneficial effects of affirmation in health information settings (see Ferrer & Cohen, 2018; Harris, 2011; Harris & Epton, 2009; 2010; McQueen & Klein, 2006; Sherman & Hartson, 2011, for reviews; see also Zhao, Peterson, Kim, & Rolfe-Redding, 2012). Two meta-analyses were published in *Health Psychology* in 2015. The first indicated aggregate effect sizes for intentions and behaviors that were significant but small in magnitude (d+ = .26, 95% confidence interval [CI] = .04 to .48; Sweeney & Moyer, 2015). The second found that “Across 34 tests of message acceptance (N = 3,433), 64 tests of intentions (N = 5,564), and 46 tests of behavior (N = 2,715), random effects models indicated small but reliable positive effects of self-affirmation on each outcome: acceptance, d+ = .17 (CI = .03 to .31); intentions, d+ = .14 (CI = .05 to .23); behavior, d+ = .32 (CI = .19 to .44) (Epton, Harris, Kane, van Koningsbruggen, & Sheeran, 2015, p. 187).” Thus, there is a wide body of supportive evidence for the efficacy of self-affirmations in facilitating adaptive health behavior in the context of threatening health information, typically in laboratory studies where information is presented in written or video form. Self-affirmation has also been extended into the field and medical settings, with beneficial effects observed as part of integrated behavioral regimen interventions with patients with chronic health conditions (e.g., Ogedegbe et al. 2012; see Cohen & Sherman, 2014, for review).

In terms of mechanisms underlying these effects or health and health-information processing, several important advances have been made. First, affirmations seem to help people calibrate to an appropriate level of threat, based on their individual risk factors, such that people with high risk can learn from information but people with low risk are not unduly alarmed (Griffin & Harris, 2011). Second, affirmations lead to greater message scrutiny (Klein, Harris, Ferrer, & Zajac, 2011; see also Binning, Brick, Cohen, & Sherman, 2015; Correll, Spencer, & Zanna, 2004). This leads to
greater openness when affirmed, but only in response to strong and persuasive rather than weak evidence. Third, affirmations lead people to focus on and attend to the high-threat content of the message that they would otherwise shun (Klein & Harris, 2009; van Koningsbruggen, Das, & Roskos-Ewoldsen, 2009). Indeed, Falk and colleagues demonstrated that affirmation changes the neural responses associated with processing self-relevant information about the dangers of a sedentary lifestyle for health, and that these neural responses predict subsequent increases in activity levels for sedentary adults (Falk et al., 2015).

Affirmations can serve as catalysts that unleash the impact of strong persuasive appeals that may otherwise be defensively rejected (Cohen & Sherman, 2014). To facilitate adaptive health behaviors, an important issue is the integration of affirmations with other interventions aimed at reducing defensiveness (such as the clinical technique of motivational interviewing; Ehret, LaBrie, Santerre, & Sherman, 2015) or other techniques such as implementation intentions, *If . . . Then . . .* plans that help people specify when, where, and how they will change their behavior (Gollwitzer & Sheeran, 2006). In the case of implementation intentions, conflicting research has been reported, with one study finding affirmation and implementation intentions leading to healthier diets up to 30 days after the study (Harris et al., 2014) and another study finding that combining affirmation and implementation intentions led to decreased exercise behavior (Jessop, Sparks, Buckland, Harris, & Churchill, 2014). A recent study sought to build on this prior work by integrating self-affirmation and implementation intentions to reduce alcohol consumption (Ehret & Sherman, 2018). The goal was to ensure that the combined approach was flexible in regard to the drinking context by being attuned to where the participants — heavy drinking students — engaged in problematic drinking behaviors, as well as present an achievable behavioral goal. When this occurred, self-affirmation and implementation intentions were effectively integrated such that drinkers were more likely to abstain up to two weeks after the experimental session (Ehret & Sherman, 2018). Identifying how to harness the power of multiple interventions is a central topic for future research.

**Physiological Stress Responses**

Self-threat can activate an individual’s stress system (Keough & Markus, 1999). By broadening the sources of self-worth, values affirmations can reduce the evaluative stress people may feel when important aspects of the self are threatened. Studies examining both acute stressful situations (Creswell et al., 2005) and chronic naturalistic stressors support this hypothesis. In one study, college students identified their most stressful midterm examination and provided urine samples to assess catecholamine levels, an indicator of sympathetic nervous system activation. Compared to baseline, students in the control condition showed an increase in epinephrine levels, whereas there was no change among those who completed two value-affirming activities in the weeks of studying and preparation for the exam (Sherman, Bunyan, Creswell, & Jaremka, 2009). While the nonaffirmed students showed a significant reduction in epinephrine levels (relative to baseline) and were significantly less likely to drop out compared to students in the other conditions, this pattern of results has only occurred among Latino American adults who experience the greatest benefit from the affirmation.

The stress-reducing effects of affirmation have several important consequences, such as the improvement in health outcomes (Creswell, Dutton, Klein, Hygge, & Luminet, 2013) and stress levels and completed academic conditions in front of an evaluation (Sherman et al., 2013). In some cases, this has led to improved performance among those who are experiencing academic stress.

**Academic Underperformance**

Educational achievement problems, such as poverty, discrimination, and differential psychological factors related to one’s social group (e.g., gender, ethnicity) or the social environment (e.g., poor academic environment, discrimination, poverty, discrimination) can help focus students on the perceived sources of stress in their lives rather than on the tasks themselves, as these tasks. In a context where students are experiencing academic stress.

Social psychological interventions have yielded long-term benefits in educational outcomes (e.g., improvement in grades, effects among middle school students who are at risk for poor academic performance). Latino American middle school students who read affirmations had significantly improved performance in their nonaffirmed Latino American peers (Sherman et al., 2013). In other cases, these interventions led to improved performance among students who are experiencing academic stress.
Strong and persuasive rather than weak and convincing. In this case, a focus on and attend to the individual’s choice of shun (Klein & Harris, 1999). Indeed, Falk and colleagues (2005) found that other responses associated with a sedentary lifestyle for stress-reducing effects. An increase in activity levels may lead to the impact of strong persuasive evidence (Ditton & Sherman, 2014). To understand how the integration of affirmation (Harris et al., 2015) may be more effective. Then, plans that help participants’ physical and behavior (Gollwitzer & Brandstatter, 1997). In some conflicting research has shown that implementation intentions led to better outcomes (Harris, & Churchill, 2014). Integrating self-affirmation into a previously effective (Ehret & Sherman, 2018) which was flexible in regard to particpants—heavy drinkers, as well as present an identity threat. Affirmation and implementation intentions were more likely to be followed (Ehret & Sherman, 2018). Academic underperformance is a central topic for educators.

Self-affirmation: Understanding the Effects (Creswell, & Jaremka, 2009). Students who were most concerned about negative evaluation in college showed the most pronounced increases in epinephrine levels (relative to baseline) and were the individuals most buffered by the affirmation. This pattern of results has occurred in many studies (e.g., Harris & Napper, 2005; Jaremka, Bunyan, Collins, & Sherman, 2011; Sherman et al., 2000, Study 1), wherein those who experience the greatest threat in a domain are the ones who experience the greatest benefit from the affirmation.

The stress-reducing effects of affirmation can lead to important downstream consequences, such as the improvement of academic performance. In one study (Creswell, Dutcher, Klein, Harris, & Levine, 2013), students indicated their chronic stress levels and completed a difficult problem-solving exercise under stressful conditions in front of an evaluator and under time pressure. Affirmation improved the performance among those who were chronically stressed, a finding with implications for those who are experiencing the stress of identity threat.

**Academic Underperformance Under Identity Threat**

Educational achievement gaps are determined by structural factors such as poverty, discrimination, and differential allocation of educational resources. An additional psychological factor that plays a role is the stress of identity threat when one’s social group (e.g., gender, racial, or ethnic group) is devalued in the academic environment (Steele, 1997). Identity threat can serve as a barrier to success in education for individuals from certain groups, such as African Americans or Latinos in the United States education system, or women in quantitative fields (Steele, 2010). However, values affirmations can make any one stressor, such as concerns about being stereotyped, less psychologically disruptive by broadening the perceived sources of self-integrity (Garcia & Cohen, 2012). This, in turn, can help focus students on the academic tasks at hand — studying, learning, taking tests — rather than on the self-evaluative implications of success or failure at these tasks. In a context where the resources for learning are available, this could improve performance.

Social psychological interventions that have featured value-affirming activities have yielded long-term benefits on both academic performance and learning (Cohen et al., 2006; 2009; Cook et al., 2012; Taylor & Walton, 2011). African American middle school students who completed in-class value-affirmation activities had improved grades, effects that continued over a two-year period (Cohen et al., 2009). Latino American middle school students who completed in-class values-affirmations had significantly improved GPA over the school year, as the affirmation deflected the downward trajectory in performance that was observed among the nonaffirmed Latino American students, with effects persisting for three years (Sherman et al., 2013). In one variation on the original procedure, teachers had access to the essays and could potentially offer more affirming feedback; this led to improved performance among African American middle school students relative
to the control group or standard affirmation (Bowen, Wegman, & Webber, 2013). Such values-affirmation interventions led to lasting effects by serving as catalysts for change, initiating positive feedback loops between the self and the social system that carried the intervention benefits forward (Cohen & Sherman, 2014). For instance, students who performed better were less likely to be assigned to the remedial track (Goyer et al., 2017), a path that can close doors of opportunity and deepen feelings of stigmatization (Steele, 1997).

Since the initial studies examining affirmation interventions to combat identity threat conducted in mixed-race middle schools (Cohen et al., 2006; 2009), there have been a number of replications and extensions among researchers and practitioners across a wide range of educational settings (e.g., Bowen et al., 2013). The effects have been successfully replicated in two studies conducted with Latino American and European American middle school students (Sherman et al., 2013), with effects not only on grades, but also on psychological outcomes (described below) such as construal level. One research team scaled up the intervention to eleven middle schools within a school district in Wisconsin and found significant effects for grade point average among the district’s ethnic and racial minority students (Borman, Grigg, & Hanselman, 2016; see also Hanselman, Rozek, Grigg, & Borman, 2017).

Contextual moderators at the level of school have been identified as well, with larger effects observed for schools that are characterized as having more threatening contexts for minorities (Hanselman, Bruch, Gamoran, & Borman, 2014; see also Dee, 2015, for another large-scale study that found no overall main effects, but contextual moderation—in particular, classrooms that provided students with more opportunities for cognitive growth yielded the predicted affirmation effect). Also supporting the important idea of contextual moderation of affirmation effects are investigations that found null effects in a wealthy suburban high school and an inner-city high school (Protzko & Aronson, 2016) and in a pre-vocational school made up largely of ethnic minority children with immigrant backgrounds in the Netherlands (de Jong, E., Jellesma, Koomen, & de Jong, P., 2016); this latter study raised the possibility that values that may be affirming in one context, such as religion that may be a potent source of affirmation in the United States, could be threatening in a European context where the immigrant youth, many of whom were Muslim and wrote about Islam, experience negative stereotypes because of their religion.

A study conducted in one mixed-ethnicity middle school had students affirm themselves up to nine times over a three-year period and observed a main effect such that affirmation benefitted all students across a range of academic attitudes and behaviors, including perceptions of the school climate, grades, attendance, and discipline (Binning et al., 2018). However, the timing of the benefit varied across gender and ethnic groups, suggesting that repeated affirmations can accumulate in the background and potentially influence a wider range of behaviors within the school ecology. Finally, cascading effects of self-affirmation have been observed in analyses of the original studies, as the more self-affirmed African Americans were within a given classroom, the better everyone’s grades were, regardless of race or condition assignment (Powers et al., 2016).

Among young adults in recent years, affirmations have found beneficial effects for threatened students. The effects have been attenuated in research with older participants, though significant improvement and gains were still seen in a range (Miyake et al., 2010). Using the principles of affirmation interventions (see also Zanna, 2015), helping them to think about an underrepresented group in a way that aligns with students historically underrepresented by affirmation intervention attenuated the effects of this strategy. This affirmation-intervention strategy was designed to increase the performance of historically underrepresented students in a gateway course by 20 percent. A comparison group of college students had a GPA of 0.21 before and after the intervention (Tibbetts et al., 2017). These effects may persist over time, and students found that affirmation acts genially (Perry et al., 2012). Moreover, found evidence that affirmed participants felt more in control of stressful environments. Among MBA students who needed self-affirmation, the values affirming (Kinias et al., 2015) merely that the affirmed personal, but not necessarily their performance in a different domain may be motivated to affirm others. Thus, across different domains, affirmations can have attenuated identity-related effects. The effects in many studies occur when the relationship between the self-system and the external world (Cohen & Sherman, 2014), social affirmation reduced self-doubt, and performance may have been affirmed (Goyer et al., 2017). Thus, in work following up the original studies (B erroneously, 2009), the value-affirmed students showed gains after the intervention (Goyer et al., 2017).

UNDERSTANDING THE TAKE-HOME MESSAGE

Across a wide variety of the domains, affirmations attenuated deficits. These effects facilitated academic performance, and the effects of the social-psychological outcomes were changes. What drives these outcomes? And the “magical” (cf., Yeager & Dweck, 2012) effects in work following up the original studies (Bowen et al., 2009), the value-affirmed students showed gains after the intervention (Goyer et al., 2017).
Among young adults in college and graduate school, studies in diverse settings have found beneficial effects of self-affirmation interventions among identity-threatened students. The gender achievement gap in college science has been attenuated in research with physics majors, with female physics majors showing significant improvement and their modal grades elevating from the C range to the B range (Miyake et al., 2010). Female engineering also benefited from being trained in the principles of affirmation in a novel approach (Walton, Logel, Peach, Spencer, & Zanna, 2015), helping them cope with the “chilly climate” of being a member of an underrepresented group in an important academic domain. First-generation college students historically underperform in pre-med courses such as biology, and an affirmation intervention attenuated this achievement gap (Harackiewicz et al., 2014).

This affirmation-intervention study showed increased retention in a crucial pre-med gateway course by 20 percent, and follow-up research shows that the first-generation college students had a GPA .18 points higher when they were affirmed three years after the intervention (Tibbitts et al., 2016). In an important study suggesting how these effects may persist over time, research with Latino American and white college students found that affirmation improved the GPA of Latino students over two years, and moreover, found evidence that this occurred because the identity-threatened students learned how to affirm themselves spontaneously when confronting serious academic stressors, like end-of-semester deadlines (Brady et al., 2016). Finally, the gender gap among MBA students was significantly attenuated among those who completed a values affirmation (Kiniias & Sim, 2016). More important, this occurred when they affirmed personal, but not organizational values, supporting the idea that affirming in a different domain may be more beneficial because it broadens the sources of self-worth from the threat. Thus, across a wide range of educational settings, values affirmations have attenuated identity-related achievement gaps. These long-term effects observed in many studies occur when they instigate a cycle of adaptive potential, an interaction between the self-system and social system that yields adaptive outcomes over time (Cohen & Sherman, 2014). For example, in the study with MBA students, the affirmation reduced self-doubt, which improved performance, and that improved performance may have been affirming and self-perpetuating (Kiniias & Sim, 2016). Indeed, in work following up the original cohort of African-American students (Cohen et al., 2009), the value-affirmed students were more likely to enroll in college many years after the intervention (Goyer et al., 2017; for discussion, see Borman, 2017).

UNDERSTANDING THE EFFECTS OF SELF-AFFIRMATIONS

Across a wide variety of threats and stressful, identity-threatening situations, values affirmations attenuated defensiveness, reduced physiological stress responses, and facilitated academic performance among students experiencing identity threats. The effects of the social-psychological manipulations fostered, in some cases, lasting changes. What drives these long-term effects when they occur? The effects are not "magical" (cf., Yeager & Walton, 2011; Wilson, 2011) but rather work through shifts in the way people construe and engage with their social environment. With
self-integrity concerns assuaged, other positive forces whose influence was suppressed by self-threat — educational, social, and persuasive — were able to exert their fuller impact (Cohen & Sherman, 2014).

Value-affirming activities encourage people to reflect on and express important, core aspects of the self. The first proposition, then, is that values affirmations boost self-resources, that is, the psychological resources that people have to cope with threats. An important part of the coping process is a determination of whether an individual has resources to cope with a focal threat or stressor (Lazarus & Folkman, 1984). Yet, as has been documented by many researchers (e.g., Pratto & John, 1991; S. E. Taylor, 1991), negative or stressful events (e.g., an important and difficult exam) tend to dominate an individual’s focus and attention, interfering with the ability to draw on one’s full range of adaptive resources. Values affirmations, by contrast, introduce a psychological resource in the form of a valued self-domain that may have helped the individual to cope with similar events in the past, and thus could conceivably be drawn upon to deal with an ongoing threat. For example, one seventh grade student in a values-affirmation study wrote: “These things are important to me because I really like playing sports with my friends a lot. Also I like being with my family and friends because I don’t want to lose them some day. Finally, I like living in the moment because I want to enjoy my life as much as I can.” In this case, the student is affirming a narrative of himself as a person who enjoys life, athletics, family and friends, a narrative he will presumably be able to sustain regardless of what may happen with a given test or on a stressful day at school.

This self-narrative as a person replete with psychological resources, strengths, and values may help people self-regulate at times when their resources would otherwise be depleted (Muraven & Baumeister, 2000). Enhanced psychological resources are inferred based on: (1) the fact that affirmations cause people to write about or reflect upon their values, relationships, and experiences, and thus these resources are likely to be salient, and (2) the finding that affirmation manipulations can counteract the effects of psychological resource depletions in a manner consistent with the strength model of self-control (Baumeister, Vohs, & Tice, 2007). It is important to recognize that the inferences about self-resources are indirect in that there is no direct measurement of psychological resources.1 While acknowledging the limitations in understanding the basic nature of psychological resources, the research suggests that affirmations can boost self-resources.

In a series of studies, laboratory induced ego-depletion was counteracted by values affirmations (Schmeichel & Vohs, 2009). In one illustrative experiment, participants watched a video while being instructed to not attend to words written on the screen, a standard ego-depleting task (see Vohs & Faber, 2007). Participants completed a values affirmation, a control activity, or a positive mood induction (to examine whether positive mood produces analogous effects as the affirmation) before being given a tedious, depleting task to perform. The prediction was that the prior depleting task

---

1 Another issue is whether affirmation increases actual self-resources or the perception of self-resources (see Clarkson, Hirt, Jia, & Alexander, 2010; Job, Dweck, & Walton, 2010).
values affirmations boost self-resources in ways that people have to cope with threatening stimuli, although the determination of whether an individual will be successful or not (Lazarus & Folkman, 1984; Pratto & John, 1991; Scrafford, 1993). Interfering with the ability to engage in self-assertions, by contrast, introduces a self-domain that may have been different in the past, and thus could constrain the self now.

For example, one seventh grade student wrote, "The things are important to me that are important to me. Also I like being with my friends and having fun all day. Finally, I like living and loving all of these things. So I can." In this case, the student’s values affirmations maintain the self regardless of what may be going on.

Psychological resources, strengths, and vulnerabilities in their resources would otherwise be dimmed. Thus, these psychological resources shape how people are motivated and thus these resources are necessary (Pyszczynski, 1988). The affirmations can counteract a threat to the self in a manner consistent with the self-affirmation theory (Walter, 2000). It is important to note that there is no direct evidence of the limitations in the self-system. The research suggests that self-depletion was counteracted by values affirmations. In this experiment, participants were asked to write about events written on the screen, a manipulation that led to the depletion of the prior depleting task. Participants then felt less threat to the perception of self-resources.

would impair persistence on this new task, unless participants were affirmed, and thus equipped with additional resources. The study supported the hypothesis, as affirmation counteracted depletion and increased persistence whereas control or positive mood inductions did not (Schmeichel & Vohs, 2009; see also Burson, Crocker, & Mischkowski, 2012).

The proposition that affirmation boosts self-resources has wide-ranging implications for outcomes such as health behaviors in which the short-term desires to eat, drink, or smoke can overwhelm the longer-term health goals. One study (Logel & Cohen, 2012) examined college-age women, the majority of whom were overweight or reported dissatisfaction with their weight and thus may have been in a chronic state of depleted resources in weight-relevant situations (Polivy & Herman, 2002; Ward & Mann, 2000). The women completed a values affirmation and were weighed on a scale, a potentially self-threatening event. Approximately 2.5 months later, they returned to the lab, weighed themselves again, and completed a test of their working memory. Participants in the affirmation condition had significantly smaller waistlines and lower body mass index at follow-up. Those in the affirmation condition also had greater working memory and for those in the affirmation condition, their increased working memory ability was associated with greater weight loss, suggesting that the affirmation increased their ability to deploy their self-regulatory resources to meet their weight-related goals (Logel & Cohen, 2012). Collectively, demonstrations that affirmations can facilitate self-regulation across a wide variety of domains suggest that they boost self-resources.

The second proposition of the theoretical account to explain the effects of values affirmations is that affirmation broadens the perspective with which people view information and events in their lives. Broader perspective here refers to a more expansive view of the self, less focused on and consumed by the threat (Cohen & Sherman, 2014). Affirmations facilitate a broader perspective by reminding people of other aspects of the self—psychological resources—that are important to them. They may also remind people of external resources, people, and relationships whom they care about beyond the threat (Crocker et al., 2008). Affirmation does not merely add additional resources in some sort of internal equation, but enables people to view a threat in a quite different manner, with greater perspective and in the context of sources of self-integrity that are not contingent upon the threatened domain.

Consistent with this, the threat of failure does not loom as large for those who complete self-affirmations. Affirmations led people to ruminate less over failures that occur in the laboratory (Koole, Smeets, van Knippenberg, & Dijksterhuis, 1999) and students to report dwelling less on what would happen if they failed during a highly stressful examination (Sherman et al., 2009). Affirmations remind people of the whole self, rather than the narrow self that might be under attack, by expanding and making salient the bases of self-worth at times of threat (Critcher & Dunning, 2015).

Affirmation-induced broader perspective can manifest itself as an overall ability to see events at a higher level of construal, to be able to appreciate the forest and not dwell on the trees. Research that has examined affirmation and object construal (Wakslak & Trope, 2009) found, across several studies, that people who were
affirmed were more likely to view objects and events at a higher level of construal. For example, they were more likely to identify locking a door in terms of the goal of the act—securing the house—than the means through which it is achieved—turning a key in the lock (Vallacher & Wegner, 1989; Waksler & Trope, 2009; Schmeichel & Vohs, 2009). However, these effects of affirmation occurred absent any clear psychological threat (cf. Brî̂nl et al., 2007), and thus provide only suggestive evidence for how affirmation affects people under threat.

When people experience threat, such as occurs for minorities who are stereotyped in educational settings (Steele, 1997), they tend to become more vigilant (Cohen & Garcia, 2008; Murphy, Steele, & Gross, 2007) on aspects of a situation that could indicate that they are at risk of being judged as a member of a stereotyped group (Steele, 2010). We hypothesized that being under a potential identity threat could lead minority students to have a lower level of construal—to focus more on detailed and concrete aspects of the situation. By contrast, self-affirmation may allow students to pull back and view events from a more relaxed and broader level of construal. We tested this in a year-long study in a mixed-ethnicity middle school composed predominantly of Latino American and white students (Sherman et al., 2013).

At multiple points during the year, students completed scales (Vallacher & Wegner, 1989) to assess their construal levels. We observed an interaction effect, such that Latino American students saw events at a higher level of construal when they were affirmed than when they were not, whereas the affirmation had no effect on the white students, the unthreatened group. Supporting the notion that identity threat could lead to a focus on concrete aspects of the situation, in the control condition, Latino American students displayed a marginal tendency to see events at a lower level of construal than white students. Taken together, these findings suggest that affirmations may help people under threat broaden the perspective with which they view events in their lives.

The third and final proposition is that affirmation leads to an uncoupling of the self and threat, reducing threat’s impact on the self (Sherman & Hartson, 2011). In the absence of affirmation, people’s self-evaluation can become engulfed by a focal threat—but with the increased breadth of perspective offered by values affirmation, focal threats can be evaluated on their own terms, with fewer self-evaluative implications.

In the realm of defensiveness, this can be operationalized as correlations between self- and threat-related variables. For example, in the political arena, individuals’ patriotism predicted their responses to an article linking U.S. foreign policy with 9/11, such that strongly patriotic people were resistant to this information, whereas nonpatriotic people were more open to it. Those at the extreme ends of the patriotism continuum were potentially the most threatened by the article, and affirmation led them to evaluate the information irrespective of their self-identified patriotism (Cohen et al., 2007), reducing political polarization (see also Sherman, Brookfield, & Ortosky, 2017). In the emotionally charged domain of sports, athletes were more self-serving and group-serving in their attributions after victory than defeat; affirmation not only reduced these biased judgments but also attenuated the correlation between threat and self-objectification of the self (Sherman & Kim, 2011), lending support to their evaluations of the threat and self-evaluation.

The realm of chronic identity threat, however, offers direct support for the decoupling of threat and self-evaluation. The subjective construal of events became associated with changes in participants as a function of self-affirmation (Cohen, 2007). Intervention across the school year substantially reduced threat and improved self-esteem for Latino American and affirmed Latino American students (Sherman et al., 2012; Sherman et al., 2013). Latino American students experienced greater threat throughout the school year compared to white students, and this threat was associated with Latino American students feeling like they did not belong or motivate, and with unaffirmed Latino American students feeling like belonging or motivation, and subjective identity threats (e.g., “I feel like I don’t belong,” identity of identity threat (e.g., “Today I don’t feel like me based on my race”). Unaffirmed Latino American students seemed to be subjectively significant rather than negatively significant traits, and their greater identity threat, their higher levels of depression, and feeling as though they belonged were all of a higher level compared to affirmed Latino American students. In both cases, the affirmation of values may have contributed to a broader evaluation of daily ups and downs that are associated with being a member of their group, and without the need for self-objectification (Sherman et al., 2013).

In a longitudinal examination of self-affirmation and academic performance, Cook and colleagues (2012) found that self-affirmation was associated with values affirmation. In this study, self-affirmation was found to be decoupled from academic self-evaluations. Participants, performance-impaired as a group, performed better in school, such that they felt better about their performance than poorly. White students’ academic performance was not related to their values affirmation. Affirmation helped reduce the uncertainty among African American students about their performance, and the impact of their performance on their self-evaluation. In the two studies (Cook et al., 2012; Sherman & Hartson, 2011), self-affirmation enabled minority students to feel better about their performance, such that they were not as negatively affected by threat and self-objectification.
correlation between them such that people evaluated their team independently of the self (Sherman & Kim, 2005). In both cases, people were less group-serving and their evaluations of the threat were more weakly associated—or decoupled—from self-evaluation.

The realm of chronic identity threat and academic performance provides more direct support for the decoupling hypothesis by examining how affirmation affects the subjective construal of minority students. Such studies test whether daily stressful events become associated—or not—with perceptions of racial threat in the minds of participants as a function of affirmation (Sherman et al., 2013; see also Walton & Cohen, 2007). Intervention studies including multiple assessments of each student across the school year assessed decoupling on a within-person level (Cook et al., 2012; Sherman et al., 2013). The studies examined whether students who perceive greater threat throughout the year also experience corresponding decrements in belonging or motivation, and how affirmation affected this relationship. In the study with Latino American and white students described above (Sherman et al., 2013), unaffirmed Latino American students had a strong association between daily stressors and subjective identity threat. On days when they experienced greater adversity and stress (e.g., “Today I feel stressed out at school”), they also experienced greater levels of identity threat (e.g., “Today in school, I am worried that other people might judge me based on my race”). Unlike white students, stressful days for unaffirmed Latino students seemed to be subjectively experienced through the lens of their race as globally significant rather than isolated events. Moreover, on days when they experienced greater identity threat, they also experienced a decrement in academic motivation, feeling as though they belonged less at school, and with decreased academic efficacy. In both cases, the affirmation eliminated these within-person correlations, suggesting that values affirmations enabled the Latino American students to experience the same daily ups and downs that all students experience, but without linking it to evaluations of their group, and without identity threat being associated with academic demoralization (Sherman et al., 2013).

In a longitudinal examination of the effect of values affirmations on belonging and academic performance of African American and white middle school students, Cook and colleagues (2012) observed a similar within-person decoupling effect of values affirmation. In this case, affirmation led perceptions of felt belonging to be decoupled from academic performance. For unaffirmed African American students, performance impinged on the extent to which they felt they belonged in the school, such that they felt more belonging in school when they performed well than poorly. White students’ feelings of belonging were generally unrelated to their performance. Affirmation decoupled the association between performance and belonging among African Americans, such that their belonging was less dependent upon their performance, and more tightly linked to the belonging they felt initially at the school (Cook et al., 2012). Although the details of the decoupling vary across the two studies (Cook et al., 2012; Sherman et al., 2013), the general point is clear: Affirmation enabled minority students to cope with threatening circumstances such that they were not as negatively impactful to their psychological state.
QUESTIONS, CHALLENGES, AND FUTURE DIRECTIONS

The theoretical account presented here draws on recent studies derived from self-affirmation theory. It presents a suite of processes—enhanced resources, broader perspective, decoupling of self and threat— theorized to be related to the important outcomes observed in affirmation studies. It is, however, only part of the story. There are other levels of analysis to be explored for a fuller understanding of the effects of values affirmation. Two examples highlight the methodologically diverse ways researchers are examining affirmation theory. Studies examining the neural signals of the brain's error-detection system found that self-affirmed participants showed increased error-related negativity (Legault, Al-Khindi, & Inzlicht, 2012), suggesting that the affirmed brain is more oriented toward opportunities to learn. Examining the content of the essays that middle school students wrote about their values revealed that writing about social belonging is key at least at this developmental stage. Affirmed African American students who wrote more about belonging themes showed the greatest academic improvement (Shnabel, Purdie-Vaughns, Cook, Garcia, & Cohen, 2013). However, the mediation is likely very context dependent; in the study of first-generation college students described earlier, affirming independence (but not belonging) was the key element when the content of the essays was examined (Tibbets et al., 2016).

Such findings can potentially be integrated into the present model. For example, social belonging is a particularly affirming self-resource in adolescence, and the discovery of its importance shines a light on the nature of affirmational resources and their efficacy at that age (Shnabel et al., 2013). It may be that greater detection of one's errors at the neural level (Legault et al., 2012) observed among affirmed participants occurs when errors are decoupled from self-evaluation—and that this decoupling could be associated with greater learning (Taylor & Walton, 2011). Such integration awaits future research.

Questions remain as to the relationship of the processes outlined here to each other and to the outcomes identified in the earlier part of this review. Although it was proposed that enhanced resources could lead to a broadened perspective on threat, it is also plausible that when people take a broader perspective they are able to draw on a wider range of resources. This mutually reinforcing nature of the components in the model presents challenges for sequencing them into discrete stages. But is the identification of discrete stages plausible considering the ongoing and reciprocal influence likely responsible for long-term effects on behavior over time (Cohen & Sherman, 2014)? The issue is further complicated because outcomes such as improved performance and reduced stress may feed back and affect the resources people draw on and the perspective that they have on threat. When multiple aspects of the model were tested within one sample, such as the assessment of construal and decoupling (Sherman et al., 2013), they were uncorrelated with each other, and moreover, neither construal nor decoupling predicted academic performance (Cook et al., 2012; Sherman et al., 2013). Although it is logically possible that the processes are independent, it is also possible that more methodologically nuanced studies will better identify the conditions for such effects. For example, Falk et al. (2013) described earlier, the deliberate selective exposure to self-affirmations (or after) grades was related to self-esteem over time (Sherman et al., 2013; Critcher, Dunbar, & Critcher, 2014; Critcher, Dunbar, 2014; Critcher, Dunbar, & Critcher, 2014), but the time that affirmations are written about could be a key factor in the long-term effects found in this research.

There are other challenges to the idea that enrichment and self-resources; performance under threat; and openness to health information might be different mediators. For example, adolescents in self-control major (Tibbets & Cohen, 2015) seem to have no effects of values affirmation on thinking about belonging and performance linked to the outcome of self-esteem. Adolescents with higher self-affirmation seem to primarily be those with high self-esteem (Tibbets & Cohen, 2015). Thus, mediation heterogeneity on different mediators and outcomes. The research approach for different people.

Despite these challenges, the case for self-affirmation. Many laboratories, from different theoretical and methodological perspectives, and methods from constructivist, cognitive, and self-control (Schmeidler & Crocker et al., 2008), theories of cognitive and affective management theory (Cohen et al., 2014) and the emerging understanding of the psychological processes involved in self-affirmation have come from many different approaches. Over the past few years. Doing so has revealed that self-affirmation can improve attitudes toward health behaviors (Cook, 2013) as catalysts for other positive changes (Sherman et al., 2013). This may have been the examination of the efforts of writing about values of self-affirmation.
NEW DIRECTIONS

Most studies derived from self-affirmation theory have focused on self-enhanced resources, broader self-schema, and self-esteem. However, only part of the story. In fact, a fuller understanding of the self-affirmation model requires methodologically diverse studies examining the neural correlates of self-affirmation. Several methodological issues were identified in studies examining the neural correlates of self-affirmation. For example, recent research has shown that self-affirmed participants exhibit greater neural activation in regions associated with self-esteem (e.g., the prefrontal cortex) than those who are not self-affirmed (e.g., Inzlicht & Schmeichel, 2012; Kring & Schmeichel, 2011). These findings are consistent with the idea that self-affirmation enhances self-esteem and self-worth, which in turn can have positive effects on cognitive processes and performance. Moreover, these effects may be particularly pronounced in situations where individuals feel threatened or under stress, as self-affirmation can help to mitigate these negative effects by promoting a sense of self-worth and self-esteem. These findings highlight the potential for self-affirmation to have broad and positive effects on a range of cognitive and psychological processes.

There are several challenges and opportunities for future research. First, the mechanisms underlying the effects of self-affirmation need to be better understood. Specifically, researchers need to identify the specific cognitive processes and neural substrates that are involved in the self-affirmation effect. This could involve using a combination of behavioral and neuroimaging methods to examine the effects of self-affirmation on cognitive processes and neural activation. Second, the practical implications of self-affirmation need to be explored. For example, how can self-affirmation be used to enhance performance in high-stress situations? This could involve designing interventions that harness the positive effects of self-affirmation to help individuals perform better in high-stress situations. Finally, the potential for self-affirmation to be used to promote well-being needs to be explored. For example, how can self-affirmation be used to help individuals develop a more positive self-concept and self-worth? This could involve designing interventions that promote self-affirmation in everyday situations, such as in the workplace or in educational settings.
work well together. The laboratory research reveals the different levels at which self-affirmations exert psychological and biological effects. The longitudinal field research demonstrates how social psychological events, and self-affirmation processes in particular, can propagate over time, and have led to basic discoveries that would not have been possible without moving from the laboratory to the field. And together, this body of work has provided an illustration of how the momentum of the self-system can perpetuate itself.

References


Self-Affirmation: Understanding the Effects


**AUTHORS’ NOTE**

David K. Sherman, Department of Psychology, University of California, Santa Barbara, CA 93106-9560, email: dsherman@ucsb.edu

Dimitri Voisin for commenting on earlier drafts of this article should be acknowledged. E-mail: dvoisin@ucsc.edu


E-mail: dhartson@ucsb.edu


E-mail: ljaremka@ucsb.edu


E-mail: jyee@ucsb.edu


E-mail: jcohen@usc.edu


E-mail: jcohen@usc.edu


E-mail: R-core@R-project.org

**Address:**

Graduate School of Education and Information Studies, University of California, Santa Barbara, CA 93106-9560, email: david@ucsb.edu

Graduate School of Business, Stanford University, Stanford, CA 94305, email: glc@stanford.edu

Graduate School of Education and Information Studies, University of California, Santa Barbara, CA 93106-9560, email: david@ucsb.edu

Graduate School of Business, Stanford University, Stanford, CA 94305, email: glc@stanford.edu

Graduate School of Education and Information Studies, University of California, Santa Barbara, CA 93106-9560, email: david@ucsb.edu

Graduate School of Business, Stanford University, Stanford, CA 94305, email: glc@stanford.edu


**AUTHORS’ NOTE**

David K. Sherman, Department of Psychological & Brain Sciences, University of California, Santa Barbara. Geoffrey L. Cohen, Graduate School of Education, Department of Psychology, and, by courtesy, Graduate School of Business, Stanford University. The authors would like to thank Cameron Brick, Phil Ehret, Kimberly Hartson, Heejung Kim, Eric Knowles, David Nussbaum, Stephen Spencer, and Dimitri Voisin for commenting on earlier versions of this manuscript. Correspondence concerning this article should be addressed to David K. Sherman at the Department of Psychological & Brain Sciences, University of California, Santa Barbara, Santa Barbara, CA 93106–9660, email: david.sherman@psych.ucsb.edu, or to Geoffrey L. Cohen, Graduate School of Education, Department of Psychology, and (by courtesy) Graduate School of Business, Stanford University, Stanford, California 94305, email: gle@stanford.edu.