



## An affirmed self and a better apology: The effect of self-affirmation on transgressors' responses to victims<sup>☆</sup>



Karina Schumann

Department of Psychology, Stanford University, Stanford, CA 94305, USA

### HIGHLIGHTS

- Tested whether self-affirmation could promote more effective apologies.
- Affirmed transgressors included more apology elements and fewer defensive strategies.
- Affirmed transgressors were thus more likely to respond in ways that boost forgiveness.
- First empirical research to identify a method for promoting more effective apologies.
- Successful application of self-affirmation theory to interpersonal conflict resolution.

### ARTICLE INFO

#### Article history:

Received 16 April 2013

Revised 22 April 2014

Available online 9 May 2014

#### Keywords:

Apology

Self-affirmation theory

Conflict resolution

Interpersonal relationship

Defensiveness

### ABSTRACT

Comprehensive apologies are powerful tools that transgressors can use to promote reconciliation with the people they have hurt. However, because many apology elements require transgressors to admit fault, express shameful emotions and promise change, transgressors often avoid these threatening elements and instead choose to use more perfunctory apologies or even defensive strategies, such as justifications or attempts to blame the person they hurt. In two studies, I aimed to increase apology comprehensiveness and reduce defensiveness using self-affirmation. I predicted that self-affirmation would help transgressors maintain their self-integrity, consequently allowing them to offer more comprehensive apologies and bypass defensive strategies. Participants received a values affirmation, recalled an unresolved conflict, and indicated what they would say to the person they had hurt. As predicted, affirmed participants offered more comprehensive apologies and used fewer defensive strategies than control participants. These studies thus identify a simple method for promoting responses that facilitate conflict resolution and demonstrate the successful application of self-affirmation to the domain of interpersonal conflict.

© 2014 Elsevier Inc. All rights reserved.

### Introduction

One of the unfortunate certainties of life is that we sometimes hurt people we care about. Luckily, these conflict events do not have to be detrimental to our relationships. Our relationship partners can forgive us for our harmful actions, and this forgiveness can increase their feelings of closeness (McCullough et al., 1998) and their willingness to cooperate and prioritize the needs of the relationship (Karremans & Van Lange, 2004). Moreover, actively discussing and working to resolve relationship problems are associated with positive feelings between partners, as well as both short- and long-term benefits to the relationship

(Overall, Sibley, & Travaglia, 2010). Thus, when managed well, conflicts can be functional and contribute to positive relationship outcomes (Gottman & Krokoff, 1989; Markman, Floyd, Stanley, & Storaasli, 1988).

When managed poorly, however, conflicts can be detrimental to relationship satisfaction, causing lasting resentment and even relationship dissolution (Carrere & Gottman, 1999; Cramer, 2000). These negative effects are not limited to romantic partnerships. Ongoing conflicts can harm other types of relationships (e.g., friendships: Raffaelli, 1997; family: Overall et al., 2010) and have consequences that extend beyond relationship outcomes. For example, unresolved conflict with a colleague in the workplace is associated with reduced organizational commitment, increased intentions to quit, and poor task performance (De Dreu & Weingart, 2003; Morrison, 2008). The ability to successfully manage and resolve interpersonal conflict thus has diverse implications for the discordant relationship, its individual members, and others in the broader social or work network.

<sup>☆</sup> This article was prepared with the support of a Social Sciences and Humanities Research Council of Canada Postdoctoral Fellowship (award # 756-2011-0397) awarded to Karina Schumann.

E-mail address: [karinas@stanford.edu](mailto:karinas@stanford.edu).

## Comprehensive apologies as tools for conflict resolution

In attempting to manage a conflict, the offending person (transgressor) can perform actions that influence whether the offended person (victim) will respond with forgiveness or continued anger and resentment. Research on conflict management suggests that an apology is one of the most powerful tools transgressors can use to promote reconciliation with the victim (Fehr, Gelfand, & Nag, 2010). Apologies increase victim forgiveness, reduce anger and aggression toward the transgressor, and validate the perceptions of the victim (e.g., Darby & Schlenker, 1982; Eaton, 2006; Exline, DeShea, & Holesman, 2007; McCullough et al., 1998; Ohbuchi, Kameda, & Agarie, 1989).

But all apologies are not created equal. Past research exploring the effects of apology composition has revealed that comprehensive apologies—those that include more basic elements of an apology—are substantially more effective at increasing victim forgiveness and decreasing blame and anger toward the transgressor (Darby & Schlenker, 1982; Kirchhoff, Wagner, & Strack, 2012; Scher & Darley, 1997; Schumann, 2012). Although the exact number of apology elements varies across frameworks proposed by different researchers, nearly all frameworks include expression of remorse, acceptance of responsibility, and offer of repair as important apology elements (Anderson, Linden, & Habra, 2006; Holmes, 1990; Kirchhoff et al., 2012; Lazare, 2004; Scher & Darley, 1997; Schonbach, 1980; Schmitt, Gollwitzer, Forster, & Montada, 2004; Schumann & Ross, 2010). In addition to these three ‘core’ elements, five other elements have been included in apology frameworks with greater variability: explanation, acknowledgement of harm, admission of wrongdoing, forbearance (a promise to behave better), and request for forgiveness (see Table 1 for a description and example of each element).

Each of these eight elements can be meaningful. For example, an offer of repair can help substantiate the apology (Minow, 2002), an explanation can help clarify the transgressor’s intentions (Lazare, 2004), and an acknowledgement of harm can validate the victim’s suffering (Eaton, 2006). By including more of these elements, transgressors can communicate a genuine attempt to take stock of their offense, repair it, and reconcile their relationship with the victim. Indeed, more comprehensive apologies appear to be more successful at promoting reconciliation (at least in part) because they are judged by victims as being more sincere—a judgment that is often needed for forgiveness to occur (Schumann, 2012; Zechmeister, Garcia, Romero, & Vas, 2004). Transgressors thus optimize their chances of being forgiven by the victim and resolving the conflict by offering more comprehensive, sincere apologies for their offenses.

## Barriers to offering comprehensive apologies

If comprehensive apologies are so effective at promoting reconciliation with the victim, why don’t transgressors use them in every conflict situation? I propose that transgressors may avoid offering comprehensive apologies because it can be threatening to do so. People are highly motivated to maintain their sense of self-worth and integrity (Sherman & Cohen, 2006), but the act of harming another person can threaten one’s identity as a good and appropriate person (Aronson, 1999; Goffman, 1971; Schlenker & Darby, 1981). Because of this threat, transgressors are likely motivated to avoid associating themselves with wrongful actions. Apology elements require transgressors to admit fault, recognize the harmful nature of their actions, promise change, convey emotions like shame or regret, and even offer a plea for forgiveness—all expressions that might diminish transgressors’ sense of power and further threaten their self-integrity (Okimoto, Wenzel, & Hedrick, 2013; Tannen, 1999, April/May). Transgressors may therefore choose to avoid using these potentially threatening elements, and instead offer more perfunctory apologies or even refuse to apologize altogether. Indeed, Okimoto et al. (2013) found that refusing to apologize boosts transgressors’ feelings of power, integrity, and state self-esteem.

Transgressors may also try to protect themselves from the negative consequences of committing an offense by responding with defensive strategies. These strategies include justifications (attempts to defend one’s behavior), victim blaming (attempts to place some or all of the responsibility for the offense on the victim), excuses (attempts to mitigate responsibility for the offense), minimizations (attempts to downplay the consequences of one’s actions), and denials (attempts to deny one’s involvement in or the presence of an offense; Itoi, Obuchi, & Fukuno, 1996; Schonbach, 1980; Scott & Lyman, 1968). Transgressors might use these defensive strategies on their own or might include them in a response that also includes apology elements (e.g., “I’m sorry [remorse] for being mean [responsibility] mom. It’s just been a long day [excuse] and you made me drive all the way from San Jose to Concord just to sleep here for a couple hours and wake up at 5 in the morning [victim blame]”). These defensive strategies can be temporarily beneficial to the transgressor by helping restore his or her self-worth, but may do so at the cost of aggravating the victim and hindering reconciliation (McLaughlin, Cody, & O’Hair, 1983; Mead, 2008). Indeed, defensiveness—refusing to take responsibility for one’s actions and instead pointing the finger of blame outward—is considered one of the most destructive behavior patterns in relationships (Gottman & Krokoff, 1989; Gottman & Silver, 1999).

Although defensive strategies may provide transgressors with short-term relief from self-integrity threat, comprehensive apologies yield

**Table 1**  
Description of apology elements and defensive strategies.

Element	Description	Example
<i>Apology elements</i>		
Remorse	Expressing a statement of apology Expressing regret or sadness about one’s actions	“I’m sorry”; “I apologize” “I feel terrible”; “I regret it”
Acceptance of responsibility	Stating that one accepts responsibility for offense Stating the offense using responsibility-accepting language	“I take full responsibility for my words” “I’m truly sorry for breaking my promise”
Repair	Offering to compensate for or fix the problem caused by one’s actions Attempting to repair the damage by making the victim feel better/loved	“I will make sure that I remember to call this week” “I love you and I am eternally grateful for all you’ve done”
Explanation	Trying to explain one’s actions without applying an external attribution	“I was afraid of commitment”
Forbearance	Promising to behave better in the future	“I’m taking steps to make sure it never happens again”
Acknowledgement of harm	Stating how the victim has suffered or been inconvenienced by one’s actions	“I know it upset you and hurt your feelings”
Admission of wrongdoing	Stating that one’s actions were wrong or unfair Stating that one should not have acted in the way that one did	“It was wrong for me to say the things I said” “I shouldn’t have spoken poorly of you”
Request for forgiveness	Asking the victim for forgiveness	“Please forgive me”
<i>Defensive strategies</i>		
Justification	Attempting to defend one’s behavior	“I’m sorry that I kicked you out, but I did it for the right reasons”
Victim blaming	Attempting to place some or all of the responsibility for the offense on the victim	“If you gave me more freedom, I wouldn’t feel the need to be dishonest”
Excuse	Attempting to mitigate responsibility for the offense	“I was very busy and in a hurry”
Minimization	Attempting to downplay the consequences of one’s actions	“I’m sorry if I upset you”; “it’s in the past”; “it was just a joke”

Note. Italicized words indicate the location of the element in the example.

positive long-term outcomes for both victims and transgressors. For example, forgiveness—the commonly demonstrated and possibly most important outcome of comprehensive apologies—enhances victims' psychological and physiological health (Lawler et al., 2005; Witvliet, Ludwig, & Vander Laan, 2001), is related to transgressors' self-forgiveness (Hall & Fincham, 2008), and ultimately boosts relationship well-being (Fincham, Beach, & Davila, 2007; Karremans & Van Lange, 2004). Given these diverse positive outcomes, I aimed to discover a method for increasing apology comprehensiveness and reducing the use of defensive strategies. Because I propose that feelings of threat pose a barrier to transgressors' willingness to offer comprehensive apologies, I examined whether self-affirmation could buffer against this threat and consequently promote more effective apologies.

### Self-affirmation as a means of promoting more effective apologies

Self-affirmation theory (Steele, 1988) posits that people can protect their self-integrity from threats by reflecting on other important values and sources of self-worth. Reflecting on core values allows people to adopt a more expansive view of the self, weakening the implications of a threat for their self-integrity. With their self-integrity intact, they can bypass defensive behaviors aimed at protecting the self from the threat (Cohen & Sherman, 2014; Sherman & Cohen, 2006). As a result, self-affirmation yields substantial benefits in a variety of domains (for a review, see Cohen & Sherman, 2014; McQueen & Klein, 2006). For example, self-affirmation increases relational security (Stinson, Logel, Shepherd, & Zanna, 2011), openness to counterattitudinal arguments (Cohen, Aronson, & Steele, 2000), and acceptance of threatening health information (Sherman, Nelson, & Steele, 2000). Past research has also demonstrated that self-affirmation can encourage openness during negotiations (Cohen et al., 2007; Ward, Atkins, Lepper, & Ross, 2011) and acknowledgement of ingroup responsibility for outgroup victimization (Cehajic-Clancy, Efron, Halperin, Liberman, & Ross, 2011), suggesting that self-affirmation might have meaningful consequences for conflict resolution.

In the research presented here, I aimed to extend past work by testing whether self-affirmation can encourage more effective responses from transgressors in the context of an interpersonal conflict. I hypothesized that giving transgressors an opportunity to affirm important values would allow them to adopt a bigger-picture focus of who they are and what is important to them. This broader perspective would put their offense in the context of a global narrative of self-integrity (Sherman, 2013), which would then allow them to focus on the needs of the victim and the relationship, rather than the need to protect their self-integrity. I therefore predicted that a self-affirmation task would help transgressors maintain their self-integrity, consequently allowing them to offer more comprehensive apologies and bypass more defensive strategies. To my knowledge, this is the first set of studies to test a method for promoting more effective apologies from transgressors.

### Study 1

In Study 1, I conducted an initial test of my hypothesis by randomly assigning adult participants to either a traditional self-affirmation condition or a control condition and then having them write what they would say to a person they had hurt. I predicted that participants who had the opportunity to affirm core values would write responses that included more apology elements and fewer defensive strategies.

### Method

#### Participants

I recruited 98 American participants (65 females, 33 males;  $M_{\text{age}} = 33.68$ ,  $SD = 12.14$ ) from Amazon Mechanical Turk (see Burhmester, Kwang, & Gosling, 2011; Paolacci, Chandler, & Ipeirotis, 2010) to

complete an online study on 'Personality and Relationships.' Two additional participants who reported conflict events from the victim's perspective were excluded from analyses.

#### Materials and procedure

Participants completed several personality measures<sup>1</sup> and then ranked 11 values and personal characteristics (e.g., creativity) in order of personal importance (Sherman et al., 2000). Participants randomly assigned to the self-affirmation condition then wrote about why their highest ranked value was important to them. Participants in the control condition wrote about why their 9th-ranked value might be important to someone else.

Next, participants were instructed to think about something they had done that had offended or hurt somebody (such as a friend, family member, romantic partner, or colleague). They were instructed to choose an offense that was currently unresolved. Participants indicated who they had offended, rated their closeness to this person (1 = *not at all close*; 7 = *extremely close*), and described the offense. They then received instructions to imagine that the person they had offended was there with them right now, and, with the intention of resolving the conflict, to write down what they would say to him or her. Participants completed demographics and a suspicion check question, and then read a debriefing letter.

Following data collection, two independent observers (blind to condition) coded the responses for each of the eight apology elements and each of the four defensive strategies<sup>2</sup> (see Table 1). Inter-observer reliability was high (average Kappa = .84); discrepancies between coders were resolved through discussion. The total number of apology elements included in each response was summed to represent apology comprehensiveness. The total number of defensive strategies included in each response was summed to represent defensive strategies.

Two observers also rated the responses for how sincerely remorseful they were (1 = *not at all*; 7 = *extremely*), and rated the offenses for severity (1 = *not at all severe*; 7 = *extremely severe*). Ratings were averaged to create indices of apology sincerity ( $r = .93$ ) and offense severity ( $r = .84$ ). Participants reported a variety of offenses that ranged broadly in severity (e.g., *low severity*: laughing at victim, forgetting to make a call; *moderate severity*: verbally attacking or insulting victim, lying; *high severity*: being unfaithful; leaving the relationship). Over a third (34.69%) of the offenses were committed against present or past romantic partners, followed by family members (30.61%), friends (20.41%), colleagues (10.20%), and acquaintances (4.08%).

### Results

As predicted, affirmed participants wrote more comprehensive apologies relative to control participants,  $t(96) = 2.32$ ,  $p = .02$ ,  $d = .47$  (see Table 2). By contrast, affirmed participants used fewer defensive strategies than did control participants,  $t(96) = -2.21$ ,  $p = .03$ ,  $d = .45$ .<sup>3</sup> Fig. 1 depicts the effect of affirmation condition on each of the elements. Although the effect of condition only reached significance or marginal significance on acceptance of responsibility, repair, and justifications, 11 of the 12 elements were influenced by self-affirmation in the predicted direction. Further, the effect of condition was significant on a

<sup>1</sup> I included three personality measures (Self-Esteem; Implicit Person Theories; Ten-Item Personality Inventory) to reduce suspicion concerning the purpose of the self-affirmation writing task. These measures did not yield any effects.

<sup>2</sup> Because participants recalled a time they had offended or hurt somebody, no participants included denials in their responses. I therefore do not include denials in the composite of defensive strategies.

<sup>3</sup> Apology comprehensiveness and defensive strategies were unassociated both in Study 1 ( $r = -.004$ ,  $p = .97$ ) and in Study 2 ( $r = .13$ ,  $p = .20$ ), which suggests that some people tend to write longer responses (including more of both types of elements) than others. In support of this interpretation, apology comprehensiveness and defensive strategies become significantly negatively correlated when statistically controlling for the number of words included in the response (Study 1:  $r = -.25$ ,  $p = .01$ ; Study 2:  $r = -.36$ ,  $p < .001$ ).

**Table 2**  
Study 1 means (and standard deviations) by affirmation condition.

	Affirmation	Control
Apology comprehensiveness	3.18 (1.97)	2.17 (2.28)
Defensive strategies	.47 (.73)	.85 (.95)
Three 'core' apology elements	1.80 (1.20)	1.17 (1.19)
Sincerely remorseful response	3.86 (1.37)	3.12 (1.59)
Offense severity	3.92 (1.09)	3.76 (1.36)
Closeness to victim	5.29 (2.04)	4.77 (2.22)

composite of the three 'core' elements (remorse, responsibility, repair) of an apology, with affirmed participants using more of these elements relative to control participants,  $t(96) = 2.61, p = .01, d = .53$ .

A significant effect of affirmation condition also emerged on observer's ratings of how sincerely remorseful the responses were, with responses offered by participants in the self-affirmation condition being rated as more sincerely remorseful than responses offered by participants in the control condition,  $t(96) = 2.42, p = .02, d = .49$ .

There was no difference between the self-affirmation and control conditions in observer-rated offense severity,  $t(96) = .63, p = .53$ , and severity was not associated with apology comprehensiveness or defensive strategies,  $ps > .52$ . Participants in the self-affirmation and control conditions also did not differ in participant-rated closeness to the victim,  $t(96) = 1.19, p = .24$ , but closeness was positively correlated with apology comprehensiveness ( $r = .32, p = .001$ ). The effects of self-affirmation on apology comprehensiveness and defensive strategies remained significant while controlling for closeness and severity,  $ps < .05$ . Ratings of closeness and severity did not interact with affirmation condition to predict apology comprehensiveness or defensive strategies,  $ps > .30$ . Participants in the affirmation condition therefore offered more comprehensive apologies and fewer defensive strategies regardless of the severity of their offense or how close they were to the victim.

## Discussion

Study 1 provides evidence for a positive effect of self-affirmation on transgressors' responses to victims. Relative to control participants, those who had previously been affirmed offered responses that included more apology elements and fewer defensive strategies. Observers also rated these responses as being more sincerely remorseful, suggesting that self-affirmation can promote transgressor responses that might

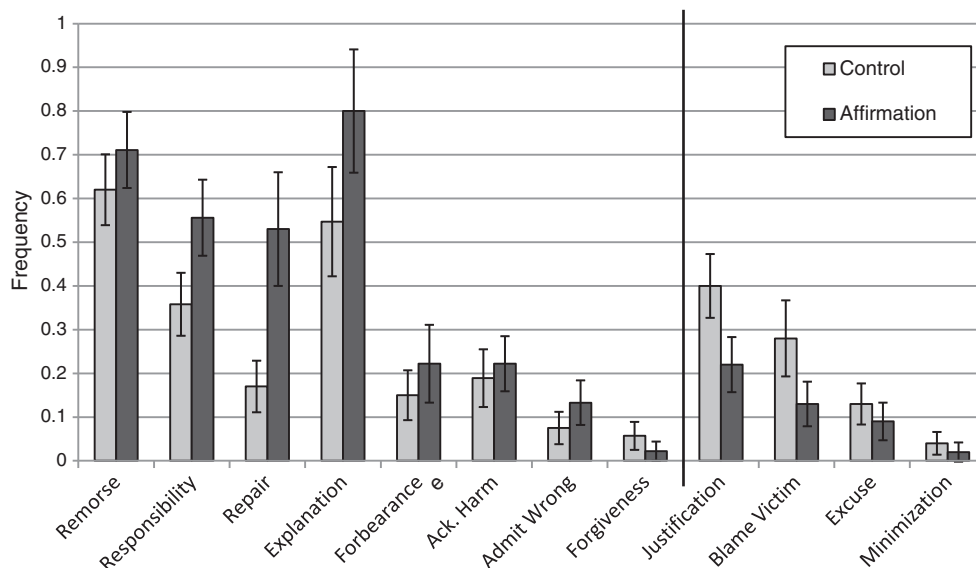
be judged as more sincere (and thus more deserving of forgiveness) by victims.

## Study 2

In Study 2, I aimed to replicate and develop a deeper understanding of the effects observed in Study 1. First, I sought to rule out mood as an alternate cause of the effect of self-affirmation on transgressors' responses. Although most evidence suggests that mood cannot explain the effects of self-affirmation (e.g., Cohen et al., 2000; Sherman et al., 2000; Shriram & Martin, 2005), some studies have shown a positive effect of self-affirmation on mood (e.g., Koole, Smeets, van Knippenberg, & Dijksterhuis, 1999), and it is plausible that transgressors experiencing more positive mood offer more comprehensive and less defensive responses.

Second, I explored mental construal level as a potential mediator. By having people write about core personal values, self-affirmation reminds them of "who they fundamentally are and what is of enduring significance to them" (Sherman et al., 2013, p. 601). This bigger-picture focus theoretically allows them to adopt a broader perspective, promoting a higher-level construal of their experiences (Schmeichel & Vohs, 2009; Sherman & Cohen, 2006; Sherman et al., 2013). In the conflict context, self-affirmation might buffer transgressors from the threat of associating themselves with an offensive behavior by promoting a broader view of the self—one that sees the offense in the context of sources of self-integrity that are not threatened—ultimately protecting their identity as a good person and allowing them to address the needs of the relationship and victim. I therefore tested whether self-affirmation—by conferring this broader perspective—would foster higher mental construal levels, which in turn would increase apology comprehensiveness and reduce defensiveness.

Third, I examined whether the timing of the affirmation has implications for its effect on transgressors' responses. Affirmations appear to be more effective when delivered before the threat, or at least before the participant has initiated a defensive response to the threat (Cohen, Garcia, Purdie-Vaughns, Apfel, & Brzustoski, 2009; Critcher, Dunning, & Armor, 2010). Self-affirmation might not undo a defensive reaction once it has been initiated, because defensive responses can also protect the self from threat. Thus, in the context of interpersonal offenses, self-affirmation might need to occur before transgressors have had an opportunity to craft a defensive response as a means of protecting their self-integrity. I tested this possibility by randomly assigning participants



**Fig. 1.** Frequency of apology elements (left) and defensive strategies (right) as a function of affirmation condition, Study 1. Error bars represent the standard error of the mean.

to either a control condition, a traditional self-affirmation condition, or a post-threat affirmation condition. To increase the chances that participants receiving the post-threat affirmation initiated defensive processing prior to the affirmation, they received foreshadowing instructions informing them that they would soon write a response to the person they had harmed (see Critcher et al., 2010).

## Method

### Participants

I recruited 96 students (78 females, 17 males, 1 missing;  $M_{\text{age}} = 24.28$ ,  $SD = 7.08$ ) from a community college to complete an online study on 'Personality and Relationships.' Three additional participants were excluded from analyses (two participants did not report an offense; 1 reported an event from the victim's perspective).

### Materials and procedure

As in Study 1, participants in the traditional affirmation/control conditions ranked 11 values and personal characteristics in order of personal importance then wrote about why their highest/9th ranked value was important to them/someone else. They then recalled a currently unresolved offense they had committed, indicated who they had offended, rated closeness and offense severity, and then described the offense.

Participants in the post-threat affirmation with foreshadowing condition completed these materials in the reverse order. They first recalled a currently unresolved offense, rated closeness and offense severity, and described the offense. Next, to foreshadow the task of writing a response to the person they had offended, participants read the following: "We are soon going to have you imagine that the person you offended was here with you right now, and ask you to write down what you would say to him or her. But before you do that, we would like you to first complete a few other measures." These foreshadowing instructions were modeled after those developed by Critcher and colleagues (2010), and were designed to initiate defensive processing prior to the affirmation. After receiving these instructions, participants completed the same affirmation materials as those in the traditional affirmation condition.

All participants then completed a state measure of mood (Brief Mood Introspection Scale [BMIS]; Mayer & Gaschke, 1988) to assess and control for any mood effects. On a 4-point scale (1 = *definitely do not feel*; 4 = *definitely feel*), participants rated their present mood using 16 adjectives (e.g., happy, grouchy, lively;  $\alpha = .89$ ). Negatively valenced items were reverse-coded so that higher scores on the composite indicated more positive mood.

Next, participants completed a shortened version of the Behavioral Identification Form (BIF; Vallacher & Wegner, 1989) to assess preferred level of mental construal. Participants saw 15 behaviors (e.g., picking an apple) and selected which one of two descriptions they preferred, one of which was low-level construal (e.g., pulling an apple off a branch), and one of which was high-level construal (e.g., getting something to eat). The number of high-level descriptions participants selected was summed to create a measure of mental construal level. Then, as in Study 1, participants received instructions to imagine that the person they had offended was there with them right now, and, with the intention of resolving the conflict, to write down what they would say to him or her.

Following data collection, two independent observers (blind to condition) coded the responses for each of the eight apology elements and four defensive strategies (average Kappa = .80); discrepancies between coders were resolved through discussion. Two additional observers also rated the responses for how sincerely remorseful they were (1 = *not at all*; 7 = *extremely*). As in Study 1, participants reported a variety of offenses that ranged broadly in severity and relationship with the victim (family members: 43.75%; present or past romantic partners: 33.33%; friends: 14.58%; colleagues: 5.21%; acquaintances: 3.13%).

## Results

A one-way ANOVA revealed a significant omnibus effect of condition on apology comprehensiveness,  $F(2, 93) = 3.90$ ,  $p = .02$  (see Table 3). As in Study 1, participants in the traditional affirmation condition wrote more comprehensive apologies relative to control participants,  $F(1, 93) = 7.23$ ,  $p = .009$ ,  $d = .71$ . Participants in the post-threat affirmation with foreshadowing condition also wrote more comprehensive apologies relative to control participants,  $F(1, 93) = 4.42$ ,  $p = .04$ ,  $d = .68$ . Participants in the traditional affirmation and post-threat affirmation conditions did not differ from each other,  $F(1, 93) = .30$ ,  $p = .59$ ,  $d = .14$ .

A one-way ANOVA on defensive strategies did not yield a significant overall omnibus effect,  $F(2, 93) = 2.17$ ,  $p = .12$ . Consistent with Study 1, however, participants in the traditional affirmation condition used marginally fewer defensive strategies relative to control participants,  $F(1, 93) = 3.46$ ,  $p = .07$ ,  $d = .50$ . Participants in the post-threat affirmation condition also used marginally fewer defensive strategies relative to control participants,  $F(1, 93) = 3.25$ ,  $p = .08$ ,  $d = .43$ . Participants in the two affirmation conditions did not differ,  $F(1, 93) = .00$ ,  $p = .99$ .

Fig. 2 depicts the effect of condition on each of the elements. Although the effect of each affirmation condition only reached significance or marginal significance on four individual elements (traditional affirmation: responsibility, acknowledge harm, request forgiveness, excuse; post-threat affirmation: responsibility, admit wrong, request forgiveness, justification), all elements were influenced by both affirmation conditions in the predicted direction. Further, relative to control participants, participants in the traditional and post-threat affirmation conditions used more of the three core apology elements,  $F(1, 93) = 4.78$ ,  $p = .03$ ,  $d = .58$ , and  $F(1, 93) = 3.19$ ,  $p = .08$ ,  $d = .48$ , respectively.

A significant effect of affirmation condition emerged on observer's ratings of how sincerely remorseful responses were,  $F(2, 93) = 3.33$ ,  $p = .04$ . Relative to responses offered by participants in the control condition, responses offered by participants in the traditional affirmation condition and post-threat affirmation condition were rated as more sincerely remorseful,  $F(1, 93) = 5.53$ ,  $p = .02$ ,  $d = .65$ , and  $F(1, 93) = 4.74$ ,  $p = .03$ ,  $d = .56$ , respectively.

There was no effect of condition on participant-rated mood,  $F(2, 93) = .10$ ,  $p = .90$ . Participants who reported more positive mood wrote responses that were marginally higher in apology comprehensiveness ( $r = .18$ ,  $p = .08$ ), but controlling for mood did not alter any condition effects,  $ps < .05$ . Mood was unassociated with defensive strategies ( $r = -.15$ ,  $p = .15$ ); controlling for mood did not alter any condition effects on defensive strategies,  $ps = .07$ . No interactions between condition and mood emerged,  $ps > .35$ .

There was also no effect of condition on mental construal level,  $F(2, 93) = 1.91$ ,  $p = .15$ . Mental construal level was unassociated with both apology comprehensiveness ( $r = .05$ ,  $p = .64$ ) and defensive strategies ( $r = .04$ ,  $p = .72$ ). Controlling for mental construal level dropped the effect of the post-threat affirmation on defensive strategies to non-significance ( $p = .11$ ), but did not alter any other condition effects. No interactions between condition and mental construal level emerged,  $ps > .21$ .

**Table 3**

Study 2 means (and standard deviations) by affirmation condition.

	Traditional affirmation	Post-threat affirmation	Control
Apology comprehensiveness	4.72 (3.82)	4.31 (3.05)	2.64 (1.68)
Defensive strategies	1.03 (1.40)	1.03 (2.04)	1.86 (1.86)
Three 'core' apology elements	2.58 (1.73)	2.44 (1.78)	1.68 (1.33)
Sincerely remorseful response	4.54 (1.38)	4.50 (1.64)	3.70 (1.21)
Positive mood	2.84 (.56)	2.79 (.58)	2.78 (.51)
Mental construal level	1.69 (.19)	1.61 (.23)	1.66 (.15)
Offense severity	4.83 (1.61)	4.13 (1.41)	3.96 (1.95)
Closeness to victim	5.39 (1.93)	5.75 (1.80)	4.82 (2.42)

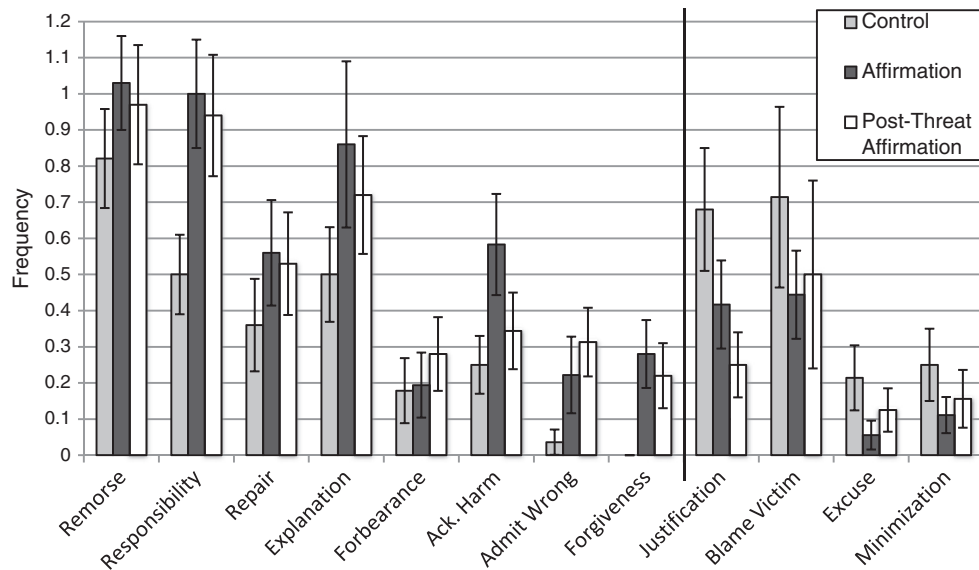


Fig. 2. Frequency of apology elements (left) and defensive strategies (right) as a function of affirmation condition, Study 2. Error bars represent the standard error of the mean.

Participants in the traditional affirmation condition rated their offenses as more severe than did control participants,  $F(1, 93) = 4.20$ ,  $p = .04$ ,  $d = .49$ . However, severity was not associated with apology comprehensiveness ( $r = .15$ ,  $p = .15$ ) or defensive strategies ( $r = -.04$ ,  $p = .74$ ), and the effects of self-affirmation remained when controlling for severity,  $p = .02$  and  $p = .06$ , respectively. Participant-rated closeness did not differ by condition, and closeness was not significantly associated with either apology comprehensiveness ( $r = .16$ ,  $p = .13$ ) or defensive strategies ( $r = -.11$ ,  $p = .28$ ). Neither ratings of severity or closeness interacted with condition to predict apology comprehensiveness or defensive strategies,  $ps > .19$ .

## Discussion

As in Study 1, affirmed participants wrote responses that included more apology elements and fewer defensives strategies relative to control participants. These effects occurred whether the affirmation was delivered before or after participants recalled a time they had harmed someone (i.e., the threat), which is inconsistent with the timing effects reported by Critcher and colleagues (2010; but see McQueen & Klein, 2006, for a meta-analysis showing no timing effects). One possible reason for this inconsistency is that defensive processing might have only been initiated when participants actually formed their responses to the person they had harmed, which took place after they had been affirmed even in the post-threat with foreshadowing condition. This finding is encouraging, as it suggests that affirmations might be effective in contexts of interpersonal conflict even if they occur after the threat of recognizing that one has committed an offense. However, more research is needed to understand timing effects and the circumstances under which affirmations may or may not be effective at facilitating conflict resolution.

Study 2 did not help clarify how self-affirmation promotes more effective responses from transgressors. As expected, there was no effect of self-affirmation on mood, and controlling for mood did not reduce the effects of self-affirmation on either apology comprehensiveness or defensive strategies. Unexpectedly, however, self-affirmation also did not influence participants' mental construal level. It is possible that using a shortened version of the BIF or presenting the BIF after participants recalled their offense or reported their mood caused this inconsistency with past research (Schmeichel & Vohs, 2009; Sherman et al., 2013), suggesting that future work on transgressors' responses might

still consider construal level as a psychological mechanism. I discuss the issue of mechanism further in the [General discussion](#) section.

One potential limitation of the methodology used in Studies 1 and 2 is that participants were instructed to select events they already perceived as offenses. This methodology does not test how self-affirmation might affect responses in more ambiguous conflict contexts. In such contexts, it is possible that self-affirmation could have detrimental effects on reconciliation by buffering the self against negative feedback and feelings of shame, consequently reducing the likelihood that transgressors recognize the offensive nature of their behavior. I therefore conducted an additional study to test whether affirmed participants are less likely than control participants to judge various ambiguous conflict behaviors as offensive.

I recruited 53 American participants (30 females, 23 males;  $M_{age} = 35.64$ ,  $SD = 14.34$ ) from Amazon Mechanical Turk to complete an online study on 'Values and Relationships.' Three additional participants who failed a question designed to catch false responding were excluded from analyses (results do not change if included). Participants were randomly assigned to complete either the self-affirmation or control materials, as in Study 1. They then completed the same measure of mood reported in Study 2 (BMIS). Next, participants imagined themselves committing 16 different behaviors against four different types of relationship partners: romantic partner, family members, colleagues, and friends. They were asked to imagine each scenario as vividly as possible, to treat each scenario as a separate event, and to try to imagine the scenarios even if they don't apply to their lives (e.g., imagining a scenario with a brother even if they don't have one). For example, they imagined that "A friend tells you an important secret. Another friend correctly guesses the secret, and asks you if it's true. Your facial expressions give it away." For each scenario, they rated the extent to which the behavior was offensive (e.g., "You have committed an offense against your friend") by indicating their agreement on a 9-point scale (1 = *strongly disagree*; 9 = *strongly agree*). Their ratings for the 16 scenarios were averaged to create an index of their tendencies to judge ambiguous conflict behaviors as offenses ( $\alpha = .83$ ).

As in Study 2, affirmation condition had no effect on positive mood,  $t(51) = -.60$ ,  $p = .55$ . Affirmation condition also had no effect on judgments of offenses,  $t(51) = .98$ ,  $p = .33$ . Participants who had been affirmed were actually non-significantly more likely to judge the behaviors as offenses ( $M = 6.10$ ,  $SD = 1.08$ ) relative to participants in the control condition ( $M = 5.80$ ,  $SD = 1.12$ ). The effect of affirmation on judgments was non-significant for each relationship type, all  $ps > .21$ .

Mood was unassociated with their judgments of offenses,  $r = -.02$ ,  $p = .92$ . Thus, although hypothetical, this additional study provides no evidence for a detrimental effect of self-affirmation in ambiguous conflict contexts.

## General discussion

Conflict is an inevitable aspect of social interactions. But if managed well, hurtful events can be transformed into constructive experiences that might even improve the relationship between the transgressor and the victim. An impressive amount of scholarship has demonstrated the benefits of forgiveness and reconciliation (e.g., Karremans & Van Lange, 2004; Witvliet et al., 2001) and has identified processes—such as comprehensive apologies—that promote reconciliation (Fehr et al., 2010). The current studies expand this literature by revealing a simple theory-based method for increasing the likelihood that transgressors will offer effective apologies. After committing an offense, transgressors' need to protect their self-integrity might prevent them from offering an appropriate response, thus further threatening the well-being of the relationships they care about. But with their self-integrity protected by self-affirmation, transgressors appear to be less likely to defend their negative behavior and more likely to apologize in a comprehensive manner that will likely encourage forgiveness. To my knowledge, this is the first empirical work to identify a method for promoting more effective responses from transgressors.

More work is needed to understand the specific psychological process underlying the effect of self-affirmation on transgressors' responses. At a higher level, self-affirmation appears to broaden people's perspectives (Sherman, 2013; Sherman & Cohen, 2006). This broader perspective could manifest itself not only as a higher construal level (Sherman et al., 2013), but also as self-transcendence (Crocker, Niiya, & Mischkowski, 2008), or greater receptivity to one's errors (Legault, Al-Khindi, & Inzlicht, 2012). Each of these potential mediators plausibly influences transgressors' responses, and should therefore be tested in future research to deepen our understanding of the effects reported here.

In addition to the question of process, other interesting questions remain. The current methodology did not permit me to code for denials or refusals to apologize, two other transgressor responses that might help transgressors protect their self-integrity but might ultimately hinder conflict resolution (Itoi et al., 1996; Okimoto et al., 2013). Because self-affirmation helps transgressors maintain their self-integrity in the face of threat, I predict that affirmed transgressors would be less likely to deny their offenses and refuse to apologize. Also, the current methodology did not assess whether participants had already attempted to reconcile with the victim but had failed. Because self-affirmation can cause goal disengagement after a salient failure experience (Vohs, Park, & Schmeichel, 2013), it is possible that self-affirmation could weaken transgressors' goals to reconcile if a recent attempt to resolve the conflict was unsuccessful. This question deserves future attention.

Finally, future work might assess whether the observed effect of self-affirmation on transgressors' responses extends beyond one conflict event. Past longitudinal work has demonstrated that self-affirmation interventions can have enduring effects by changing the trajectory of recursive processes (Cohen et al., 2009). For example, self-affirmation improved relational security up to eight weeks later, possibly by introducing a recursive loop of increases in relational security, positive social outcomes, and, consequently, additional increases in relational security (Stinson et al., 2011). Conceivably, self-affirmation also increased long-term relational security by reducing defensiveness and promoting more effective responses to negative behaviors, which ultimately bolstered the relationship. Future work might explore this possibility and examine whether self-affirmation can create a recursive process of constructive responses from both transgressors and victims, which could lead to large and long-lasting positive outcomes for the relationship and its individual members.

## References

- Anderson, J. C., Linden, W., & Habra, M. E. (2006). Influence of apologies and trait hostility on recovery from anger. *Journal of Behavioral Medicine*, 29, 347–358.
- Aronson, E. (1999). *Dissonance, hypocrisy, and the self-concept*. Readings about the social animal. New York: Worth.
- Burhmester, M., Kwang, T., & Gosling, S. D. (2011). Amazon's Mechanical Turk: A new source of inexpensive, yet high quality, data? *Perspectives on Psychological Science*, 6, 3–5.
- Carrere, S., & Gottman, J. G. (1999). Predicting divorce among newlyweds from the first three minutes of a marital conflict discussion. *Family Process*, 38, 293–301.
- Cehajic-Clancy, S., Effron, D. A., Halperin, E., Liberman, V., & Ross, L. D. (2011). Affirmation, acknowledgment of ingroup responsibility, group-based guilt, and support for reparative measures. *Journal of Personality and Social Psychology*, 101, 256–270.
- Cohen, G. L., Aronson, J., & Steele, C. M. (2000). When beliefs yield to evidence: Reducing biased evaluation by affirming the self. *Personality and Social Psychology Bulletin*, 26, 1151–1164.
- Cohen, G. L., Garcia, J., Purdie-Vaughns, V., Apfel, N., & Brzustoski, P. (2009). Recursive processes in self-affirmation: Intervening to close the minority achievement gap. *Science*, 324, 400–403.
- Cohen, G. L., & Sherman, D. K. (2014). The psychology of change: Self-affirmation and social psychological intervention. *Annual Review of Psychology*, 65, 333–371.
- 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.
- Cramer, D. (2000). Relationship satisfaction and conflict style in romantic relationships. *The Journal of Psychology*, 134, 337–341.
- 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.
- Darby, B. W., & Schlenker, B. R. (1982). Children's reactions to apologies. *Journal of Personality and Social Psychology*, 43, 742–753.
- De Dreu, C. K. W., & Weingart, L. R. (2003). Task versus relationship conflict, team performance, and team member satisfaction: A meta-analysis. *Journal of Applied Psychology*, 88, 741–749.
- Eaton, J. A. (2006). Repentance as validation: Toward an understanding of the mechanisms behind repentance and forgiveness. *Dissertation Abstracts, Section B: The Sciences and Engineering*, 67(1-B), 598.
- Exline, J. J., DeShea, L., & Holesman, V. T. (2007). Is apology worth the risk? Predictors, outcomes, and ways to avoid regret. *Journal of Social and Clinical Psychology*, 26, 479–504.
- Fehr, R., Gelfand, M. J., & Nag, M. (2010). The road to forgiveness: A meta-analytic synthesis of its situational and dispositional correlates. *Psychological Bulletin*, 136, 894–914.
- Fincham, F. D., Beach, S. R. H., & Davila, J. (2007). Longitudinal relations between forgiveness and conflict resolution in marriage. *Journal of Family Psychology*, 21, 542–545.
- Goffman, E. (1971). *Relations in public*. New York: Basic Books.
- Gottman, J. M., & Krokoff, L. J. (1989). Marital interaction and satisfaction: A longitudinal view. *Journal of Consulting and Clinical Psychology*, 57, 47–52.
- Gottman, J., & Silver, N. (1999). *The seven principles for making marriage work: A practical guide from the country's foremost relationship expert*. New York, NY: Three Rivers Press.
- Hall, J. H., & Fincham, F. D. (2008). The temporal course of self-forgiveness. *Journal of Social and Clinical Psychology*, 27, 174–202.
- Holmes, J. (1990). Apologies in New Zealand English. *Language in Society*, 19(2), 155–199.
- Itoi, R., Obuchi, K. I., & Fukuno, M. (1996). A cross-cultural study of preference of accounts: Relationship closeness, harm severity, and motives of account making. *Journal of Applied Social Psychology*, 26, 913–934.
- Karremans, J. C., & Van Lange, P. A. M. (2004). Back to caring after being hurt: The role of forgiveness. *European Journal of Social Psychology*, 34, 207–227.
- Kirchhoff, J., Wagner, U., & Strack, M. (2012). Apologies: Words of magic? The role of verbal components, anger reduction, and offense severity. *Peace and Conflict: Journal of Peace Psychology*, 18, 109–130.
- 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.
- Lawler, K. A., Younger, J. W., Piferi, R. L., Jobe, R. L., Edmondson, K. A., & Jones, W. H. (2005). The unique effects of forgiveness on health: An exploration of pathways. *Journal of Behavioral Medicine*, 28, 157–167.
- Lazare, A. (2004). *On apology*. New York: Oxford University Press.
- Legault, L., Al-Khindi, T., & Inzlicht, M. (2012). Preserving integrity in the face of performance threat: Self-affirmation enhances neurophysiological responsiveness to task errors. *Psychological Science*, 23, 1455–1460.
- Markman, H. J., Floyd, F. J., Stanley, S. M., & Storaasli, R. D. (1988). The prevention of marital distress: A longitudinal investigation. *Journal of Consulting and Clinical Psychology*, 56, 210–217.
- Mayer, J. D., & Gaschke, Y. N. (1988). The experience and meta-experience of mood. *Journal of Personality and Social Psychology*, 55, 102–111.
- McCullough, M. E., Rachal, K. C., Sandage, S. J., Worthington, E. L., Jr., Brown, S. W., & Hight, T. L. (1998). Interpersonal forgiving in close relationships: II. Theoretical elaboration and measurement. *Journal of Personality and Social Psychology*, 75, 1586–1603.
- McLaughlin, M. L., Cody, M. J., & O'Hair, H. D. (1983). The management of failure events: Some contextual determinants of accounting behavior. *Human Communication Research*, 9, 208–224.
- McQueen, A., & Klein, W. M. P. (2006). Experimental manipulation of self-affirmation: A systematic review. *Self and Identity*, 5, 289–354.

- Mead, B. G. (2008). When do we forgive? An examination of relational and apologetic factors that are influential in forgiveness with friendship relationships. *Dissertation Abstracts, Section B: The Sciences and Engineering*, 69(2-B), 1379.
- Minow, M. (2002). *Breaking the cycles of hatred*. Princeton, NJ: Princeton University Press.
- Morrison, R. L. (2008). Negative relationships in the workplace: Associations with organisational commitment, cohesion, job satisfaction, and intention to turnover. *Journal of Management and Organization*, 14, 330–344.
- Ohbuchi, K., Kameda, M., & Agarie, N. (1989). Apology as aggression control: Its role in mediating appraisal of and response to harm. *Journal of Personality and Social Psychology*, 56, 219–227.
- Okimoto, T. G., Wenzel, M., & Hedrick, K. (2013). Refusing to apologize can have psychological benefits (and we issue not mea culpa for this research finding). *European Journal of Social Psychology*, 43, 22–31.
- Overall, N. C., Sibley, C. G., & Travaglia, L. K. (2010). Loyal but ignored: The benefits and costs of constructive communication behavior. *Personal Relationships*, 17, 127–148.
- Paolacci, G., Chandler, J., & Ipeirotis, P. G. (2010). Running experiments on Amazon Mechanical Turk. *Judgment and Decision Making*, 5, 411–419.
- Raffaelli, M. (1997). Young adolescents' conflicts with siblings and friends. *Journal of Youth and Adolescence*, 26, 539–558.
- Scher, S. J., & Darley, J. M. (1997). How effective are the things people say to apologize? Effects of the realization of the apology speech act. *Journal of Psycholinguistic Research*, 26, 127–140.
- Schlenker, B. R., & Darby, B. W. (1981). The use of apologies in social predicaments. *Social Psychology Quarterly*, 44, 271–278.
- Schmitt, M., Gollwitzer, M., Forster, N., & Montada, L. (2004). Effects of objective and subjective account components on forgiving. *Journal of Social Psychology*, 144, 465–485.
- Schonbach, P. (1980). A category system for account phases. *European Journal of Social Psychology*, 10, 195–200.
- Schmeichel, B. J., & Vohs, K. D. (2009). Self affirmation and self-control: Affirming core values counteracts ego depletion. *Journal of Personality and Social Psychology*, 96, 770–782.
- Schumann, K. (2012). Does love mean never having to say you're sorry? Associations between relationship satisfaction, perceived apology sincerity, and forgiveness. *Journal of Social and Personal Relationships*, 29, 997–1010.
- Schumann, K., & Ross, M. (2010). Why women apologize more than men: Gender differences in thresholds for offensive behavior. *Psychological Science*, 21, 1649–1655.
- Scott, M. B., & Lyman, S. M. (1968). Accounts. *American Sociological Review*, 33, 46–62.
- Sherman, D. K. (2013). Self-affirmation: Understanding the effects. *Social and Personality Psychology Compass*, 7, 834–845.
- Sherman, D. K., & Cohen, G. L. (2006). The psychology of self-defense: Self-affirmation theory. *Advances in Experimental Social Psychology*, 38, 138–242.
- Sherman, D. K., Hartson, K. A., Binning, K. R., Purdie-Vaughns, V., Garcia, J., Taborsky-Barba, S., et al. (2013). Deflecting the trajectory and changing the narrative: How self-affirmation affects academic performance and motivation under identity threat. *Journal of Personality and Social Psychology*, 104, 591–618.
- Sherman, D. A. K., Nelson, L. D., & Steele, C. M. (2000). Do messages about health risks threaten the self? Increasing the acceptance of threatening health messages via self-affirmation. *Personality and Social Psychology Bulletin*, 26, 1046–1058.
- Shrira, I., & Martin, L. L. (2005). Stereotyping, self-affirmation, and the cerebral hemispheres. *Personality and Social Psychology Bulletin*, 31, 846–856.
- Steele, C. M. (1988). The psychology of self-affirmation: Sustaining the integrity of the self. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (pp. 261–302). New York: Academic Press.
- Stinson, D. A., Logel, C., Shepherd, S., & Zanna, M. P. (2011). Rewriting the self-fulfilling prophecy of social rejection: Self-affirmation improves relational security and social behavior up to 2 months later. *Psychological Science*, 22, 1145–1149.
- Tannen, D. (1999, April/May). Contribute makes right. *Civilization*, 6, 67–70.
- Vallacher, R. R., & Wegner, D.M. (1989). Levels of personal agency: Individual variation in action identification. *Journal of Personality and Social Psychology*, 57, 660–671.
- Vohs, K. D., Park, J., & Schmeichel, B. J. (2013). Self-affirmation can enable goal disengagement. *Journal of Personality and Social Psychology*, 104, 14–27.
- Ward, A., Atkins, D. C., Lepper, M. R., & Ross, L. (2011). Affirming the self to promote agreement with another: Lowering a psychological barrier to conflict resolution. *Personality and Social Psychology Bulletin*, 37, 1216–1228.
- Witvliet, C. V., Ludwig, T. E., & Vander Laan, K. L. (2001). Granting forgiveness or harboring grudges: Implications for emotion, physiology, and health. *Psychological Science*, 12, 117–123.
- Zechmeister, J. S., Garcia, S., Romero, C., & Vas, S. N. (2004). Don't apologize unless you mean it: A laboratory investigation of forgiveness and retaliation. *Journal of Social and Clinical Psychology*, 23, 532–564.