Reducing narcissistic aggression by buttressing self-esteem: An experimental field study
Thomaes, S.; Bushman, B.J.; Orobio de Castro, B.; Cohen, G.L.; Denissen, J.J.A.

published in
Psychological Science
2009

DOI (link to publisher)
10.1111/j.1467-9280.2009.02478.x

document version
Publisher’s PDF, also known as Version of record

Link to publication in VU Research Portal

citation for published version (APA)

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

E-mail address:
vuresearchportal.ub@vu.nl

Download date: 04. Mar. 2022
Reducing Narcissistic Aggression by Buttressing Self-Esteem: An Experimental Field Study
Sander Thomaes, Brad J. Bushman, Bram Orobio de Castro, Geoffrey L. Cohen and Jaap J.A. Denissen

*Psychological Science* 2009 20: 1536
DOI: 10.1111/j.1467-9280.2009.02478.x

The online version of this article can be found at:
http://pss.sagepub.com/content/20/12/1536
Reducing Narcissistic Aggression by Buttressing Self-Esteem

An Experimental Field Study

Sander Thomaes, Brad J. Bushman, Bram Orobio de Castro, Geoffrey L. Cohen, and Jaap J.A. Denissen

Department of Psychology, Utrecht University; Institute for Social Research, University of Michigan; Department of Communication Science, VU University Amsterdam; Department of Psychology, University of Colorado; and Department of Psychology, Humboldt University of Berlin

ABSTRACT—Narcissistic individuals are prone to become aggressive when their egos are threatened. We report a randomized field experiment that tested whether a social-psychological intervention designed to lessen the impact of ego threat reduces narcissistic aggression. A sample of 405 young adolescents (mean age = 13.9 years) were randomly assigned to complete either a short self-affirmation writing assignment (which allowed them to reflect on their personally important values) or a control writing assignment. We expected that the self-affirmation would temporarily attenuate the ego-protective motivations that normally drive narcissists’ aggression. As expected, the self-affirmation writing assignment reduced narcissistic aggression for a period of a school week, that is, for a period up to 400 times the duration of the intervention itself. These results provide the first empirical demonstration that buttressing self-esteem (as opposed to boosting self-esteem) can be effective at reducing aggression in at-risk youth.

Aggression in schools is a serious problem worldwide. Children are exposed to physical violence, verbal assaults, and psychological abuse at their schools on a daily basis (Kochenderfer-Ladd & Ladd, 2001; Nishina & Juvonen, 2005). Many current intervention programs rely on “boosting self-esteem” to reduce aggression (e.g., Kusche & Greenberg, 1994; Ringwalt, Graham, Paschall, Flewelling, & Browne, 1996). However, contrary to popular wisdom, aggressive people do not typically have low self-esteem. Instead, they often have grandiose, inflated, narcissistic self-views. Narcissistic individuals—both adults and children—are especially likely to lash out and become aggressive when their egos are threatened (e.g., Bushman & Baumeister, 1998; Stucke & Sporer, 2002; Thomaes, Bushman, Stegge, & Olthof, 2008). Thus, there are no compelling theoretical or empirical reasons to suggest that boosting self-esteem will be effective in reducing aggression. “Buttressing self-esteem” (i.e., making self-esteem less vulnerable to ego threat) should be more effective, at least in narcissistic individuals. Interventions aimed at buttressing self-esteem lessen the psychological impact of ego threat by focusing individuals on the core traits that define them as a person. Such interventions do not artificially raise, or inflate, self-esteem (Crocker, Niya, & Mischkowski, 2008). The study we report here tested whether a short self-affirmation writing exercise known to temporarily buttress individuals’ self-esteem can reduce narcissistic aggression.

INTERVENING WITH NARCISSISTIC AGGRESSION

Normal narcissism (i.e., narcissism viewed as a continuous trait, not a personality disorder) involves grandiose but simultaneously vulnerable self-views that are found in general child and adult populations (Raskin & Terry, 1983; Thomaes, Stegge, Bushman, Olthof, & Denissen, 2008). Research shows that narcissistic self-views are highly contingent on evaluations by others (Morf & Rhodewalt, 2001). Narcissists crave admiration and respect from others, and they are quick to engage in self-regulatory strategies to protect their self-views when they need to. Accordingly, researchers have explained narcissists’ aggressive reactions to ego threat as defensive attempts to maintain self-worth (Bushman & Baumeister, 1998; Morf & Rhodewalt, 2001).
Thus, intervention techniques able to buffer people's self-views against ego threat should reduce narcissistic aggression. One such technique is to allow individuals to reaffirm their sense of self (Sherman & Cohen, 2006; Steele, 1988). Self-affirmation theory holds that an individual's overall sense of self is based on multiple domains of functioning, and that a threat to one domain of functioning can be compensated for by reflecting on the personal importance of a different domain (such as a self-defining skill or interest). Previous research has shown that self-affirmations buttress self-esteem, and thereby reduce the psychological impact of threatening feedback and social-evaluative stress both in the laboratory (Creswell et al., 2005; Koole, Smeets, Van Knippenberg, & Dijksterhuis, 1999; Sherman & Cohen, 2002) and in actual classroom settings (Cohen, Garcia, Apfel, & Master, 2006).

THE PRESENT STUDY

The present study tested whether a self-affirmation intervention can reduce narcissistic aggression in the "real world." Participants were 12 to 15 years old. We studied children this age for four reasons. First, ego threat is more frequently experienced in early adolescence than in any other developmental period. Children this age are increasingly concerned about blows to their self-esteem (Harter, 2006; Nishina & Juvonen, 2005; Rosenberg, 1986). Second, ego threat is particularly consequential in early adolescence, because children this age—in contrast to young children—are able to make global negative evaluations of the self (e.g., "I am a worthless person") that make ego-threatening experiences potentially harmful (Ferguson, Stegge, & Damhuis, 1991). Third, the extent to which children engage in serious aggressive and violent behavior increases steeply in early adolescence (Dodge, Coie, & Lynam, 2006). Fourth, it seems important to try to intervene with individuals' self-views in a developmental period when self-views start to take a relatively mature form, but have not yet become deeply ingrained in patterns of maladaptive behavior that may be hard to change.

We conducted a randomized field experiment in which participants completed either a short (15-min) self-affirmation or a control writing exercise in their classes (Cohen et al., 2006). In the affirmation condition, participants wrote about their most important values and why these values may be important to them. In the control condition, participants wrote about their least important values and why these values may be important to other people. Peer reports of aggressive behavior in the schools served as an ecologically valid measure of aggression. We also obtained reports of state self-esteem, a continuous measure of experienced ego threat. Low state self-esteem is the key experiential component of ego threat (Baumeister, Smart, & Boden, 1996). Narcissism was measured along with trait self-esteem 3 weeks before the self-affirmation intervention. Aggression and state self-esteem were measured in the week before the self-affirmation (Assessment 1), in the week after the self-affirmation (Assessment 2), in the week after a second self-affirmation (Assessment 3), and again 3 weeks later (Assessment 4). On the basis of previous laboratory experiments (e.g., Bushman & Baumeister, 1998; Stucke & Sporer, 2002; Thomaes, Bushman, et al., 2008), we predicted that narcissistic youth would behave aggressively, but only when they reported having low state self-esteem (i.e., when they experienced high levels of ego threat). More important, we predicted that our self-affirmation intervention would reduce aggression in narcissistic youth having low state self-esteem. Our short-term longitudinal design permitted us to test the directionality of effects, and we conducted lagged-effects analyses to establish that the intervention indeed influenced narcissists' aggression after they experienced low state self-esteem (as predicted), rather than narcissists' experience of low self-esteem after they behaved aggressively.

METHOD

Participants
Participants were 405 sixth and seventh graders (52% boys, 48% girls) recruited from two public middle schools serving middle-class neighborhoods in The Netherlands (parental consent rate = 96%). They ranged in age from 12 to 15 (mean age = 13.9 years, SD = 0.7). Most participants were Caucasian (90%); 10% had other (e.g., Turkish, Dutch Antillean, mixed) cultural-ethnic backgrounds.

Self-View Measures
Three weeks before the start of the experiment, students completed self-report measures of narcissism and trait self-esteem in their classrooms. Trait self-esteem was measured to examine the possibility that low trait self-esteem contributes to real-world aggression, as has been suggested by some researchers (Donnellan, Trzesniewski, Robins, Moffitt, & Caspi, 2005). Trait self-esteem was measured using the 5-item Global Self-Worth subscale of the Self-Perception Profile for Adolescents (Harter, 1988; Cronbach’s α = .76). This reliable and valid scale assesses adolescents' overall perception of worth as a person (e.g., "Some kids like the kind of person they are. How much are you like these kids?"). Items are rated along a 4-point scale (0 = not at all, 3 = exactly). Narcissism was measured using the 10-item Childhood Narcissism Scale (Thomaes, Stegge, et al., 2008; Cronbach’s α = .77). This scale assesses grandiose, entitled views of self and adversarial interpersonal attitudes (e.g., “Without me, our class would be much less fun” and “Kids like me deserve something extra”). Items are rated along a 4-point scale (0 = not at all true, 3 = completely true). The Childhood Narcissism Scale is a reliable, one-dimensional measure of stable individual differences in childhood narcissism. Research indicates that childhood narcissism has psychological and in-
terpersonal correlates very similar to those of adult narcissism (Thomaes, Stegge, et al., 2008).

### Procedure

State self-esteem and aggression were first measured at Assessment 1, which was completed on Friday afternoon in the week prior to the first self-affirmation intervention. To measure state self-esteem, we presented students with a pictorial scale showing a very small figure at one end and a very large figure at the other end (the scale was taken from Bradley & Lang’s, 1994, Self-Assessment Manikin). The small figure was labeled “very unsatisfied with myself in the past week,” and the large figure was labeled “very satisfied with myself in the past week.” Students indicated which figure on the 9-point scale best reflected how they felt about themselves in the past week. Next, students completed a peer-nomination aggression measure developed in a pilot study (see Aggression Measure Pilot Study in the Supporting Information available on-line—see p. 1542). The measure contained 1 item for physical aggression (“Who kicked, pushed, or hit another student at school in the past week?”), 1 item for direct verbal aggression (“Who called another student names, or said mean things to another student at school in the past week?”), 1 item for relational aggression (“Who spread rumors or lies about another student, or excluded another student from the group at school in the past week?”), and 4 positively worded filler items (e.g., “Who seemed very happy in the past week?”). Students circled the names of all classmates (on a class roster with order randomized) for whom each item applied. For each student, the number of received nominations was summed across the 3 aggression items and divided by the number of classmates to yield a weekly aggression score (Cronbach’s $\alpha = .74$ at Assessment 1).

The following Monday morning, participants completed the intervention exercises in their classrooms. Each individual was randomly assigned to either the self-affirmation condition or the control condition. In each class, there were approximately equal numbers of participants in the two conditions, and the gender distribution was also approximately equal (53% boys and 47% girls in the self-affirmation condition; 52% boys and 48% girls in the control condition). Following standard procedures (Cohen et al., 2006), students were given a list of 12 values (i.e., athletic ability, being good at art, being smart or getting good grades, being creative, being independent, living in the moment, belonging to a social group, music, politics, relationships with friends or family, religious values, sense of humor). In the self-affirmation condition, students selected 2 or 3 of their most important values and then wrote a short paragraph about why these values were important to them. In the control condition, students selected 2 or 3 of their least important values and then wrote about why these values may be important to other people. To reinforce the manipulation, we also asked students to indicate their level of agreement with several statements about the values they chose (e.g., “I care about these values” in the self-affirmation condition and “Some people care about these values” in the control condition). Students worked on the exercises quietly and independently, and returned their work in a sealed envelope after they finished. The exercises took approximately 15 min to complete. Students who were not present on Monday ($n = 9; 2\%$) completed the intervention exercises the first day they reentered school.

On Friday afternoon in the same week (i.e., 1 school week after the first intervention), Assessment 2 of state self-esteem and aggression was completed. The measures were identical to the ones completed at baseline (Cronbach’s $\alpha = .76$ for Assessment 2 aggression). To keep students motivated, we held a raffle for a CD or DVD among the participants in each class.

Five weeks later on Monday morning, students completed a second intervention exercise (or “booster shot affirmation”—Cohen et al., 2006). They were assigned to the same condition to which they were assigned previously. The exercises and procedures were the same as for the first intervention. On Friday afternoon in the same week (i.e., 6 school weeks after the first intervention and 1 school week after the second), Assessment 3 of state self-esteem and aggression was completed (Cronbach’s $\alpha = .82$ for Assessment 3 aggression). Finally, 3 weeks later on Friday afternoon (i.e., 9 school weeks after the first intervention and 4 school weeks after the second), students completed Assessment 4 of state self-esteem and aggression (Cronbach’s $\alpha = .77$ for Assessment 4 aggression).

### RESULTS

#### Preliminary Analyses

Narcissism, trait self-esteem, baseline state self-esteem, baseline aggression, gender distribution, and age did not differ between groups ($ps > .39, p_{reps} < .58$). Thus, random assignment to the self-affirmation and control groups was successful. (Table S1 in the Supporting Information available on-line provides descriptive statistics and correlations.)

#### Primary Analyses

The data were analyzed using hierarchical linear modeling (SPSS mixed). They were organized to account for their hierarchical structure, with four assessment occasions (with an auto-regressed AR1 covariance structure) nested within students. We tested two models with aggression as the dependent variable. Gender and trait self-esteem were included as covariates. Narcissism and state self-esteem were included as predictor variables. In addition, three dummy variables were included as predictor variables: one indicating group assignment ($0 = $ control condition, $1 = $ self-affirmation condition), one indicating the short-term intervention effect ($0 = $ control condition 1 school week ago, $1 = $ self-affirmation condition 1 school week ago), and one indicating the long-term intervention effect ($0 = $
Narcissism was associated with increased aggression, regardless of whether students had low or high state self-esteem, \( b = 0.15, p > .22, p_{\text{rep}} < .70 \), and \( b = -0.22, p > .07, p_{\text{rep}} < .86 \), respectively (see Fig. 1). Thus, these analyses indicate that a 15-min self-affirmation writing exercise reduces narcissistic aggression for a period of 1 school week. There was no significant three-way interaction effect involving the long-term intervention dummy variable, \( p < .10, ^*p < .05, ^{*}*p < .01 \).

In Model 1, aggression was predicted by concurrent levels of state self-esteem (see Table 1). There was a significant interaction between narcissism and state self-esteem, \( b = -0.09, p < .001, p_{\text{rep}} > .98 \). More important, this two-way interaction was qualified by the predicted three-way interaction of narcissism, state self-esteem, and the short-term intervention dummy variable, \( b = 0.10, p < .02, p_{\text{rep}} > .93 \). To interpret this significant three-way interaction, we examined the two-way interactions between narcissism and state self-esteem separately for the control and self-affirmation conditions. In the control condition, the standard pattern found in previous laboratory research emerged. Narcissism was associated with increased aggression when students had a low level of state self-esteem (1 SD below the mean; Aiken & West, 1991), \( b = 0.35, p < .01, p_{\text{rep}} > .95 \), but not when students had a high level of state self-esteem (1 SD above the mean), \( b = 0.05, p > .68, p_{\text{rep}} < .37 \) (see Fig. 1). Thus, we generalized existing laboratory findings to the real world. By contrast, in the self-affirmation condition, narcissism was not associated with increased aggression, regardless of whether students had low or high state self-esteem, \( b = 0.15, p > .22, p_{\text{rep}} < .70 \), and \( b = -0.22, p > .07, p_{\text{rep}} < .86 \), respectively (see Fig. 1). Thus, these analyses indicate that a 15-min self-affirmation writing exercise reduces narcissistic aggression for a period of 1 school week. There was no significant three-way interaction effect involving the long-term intervention dummy variable, \( p < .10, ^*p < .05, ^{*}*p < .01 \).

In Model 1, aggression was predicted by concurrent levels of state self-esteem (see Table 1). There was a significant interaction between narcissism and state self-esteem, \( b = -0.09, p < .001, p_{\text{rep}} > .98 \). More important, this two-way interaction was qualified by the predicted three-way interaction of narcissism, state self-esteem, and the short-term intervention dummy variable, \( b = 0.10, p < .02, p_{\text{rep}} > .93 \). To interpret this significant three-way interaction, we examined the two-way interactions between narcissism and state self-esteem separately for the control and self-affirmation conditions. In the control condition, the standard pattern found in previous laboratory research emerged. Narcissism was associated with increased aggression when students had a low level of state self-esteem (1 SD below the mean; Aiken & West, 1991), \( b = 0.35, p < .01, p_{\text{rep}} > .95 \), but not when students had a high level of state self-esteem (1 SD above the mean), \( b = 0.05, p > .68, p_{\text{rep}} < .37 \) (see Fig. 1). Thus, we generalized existing laboratory findings to the real world. By contrast, in the self-affirmation condition, narcissism was not associated with increased aggression, regardless of whether students had low or high state self-esteem, \( b = 0.15, p > .22, p_{\text{rep}} < .70 \), and \( b = -0.22, p > .07, p_{\text{rep}} < .86 \), respectively (see Fig. 1). Thus, these analyses indicate that a 15-min self-affirmation writing exercise reduces narcissistic aggression for a period of 1 school week. There was no significant three-way interaction effect involving the long-term intervention dummy variable, \( p < .10, ^*p < .05, ^{*}*p < .01 \).

In Model 1, aggression was predicted by concurrent levels of state self-esteem (see Table 1). There was a significant interaction between narcissism and state self-esteem, \( b = -0.09, p < .001, p_{\text{rep}} > .98 \). More important, this two-way interaction was qualified by the predicted three-way interaction of narcissism, state self-esteem, and the short-term intervention dummy variable, \( b = 0.10, p < .02, p_{\text{rep}} > .93 \). To interpret this significant three-way interaction, we examined the two-way interactions between narcissism and state self-esteem separately for the control and self-affirmation conditions. In the control condition, the standard pattern found in previous laboratory research emerged. Narcissism was associated with increased aggression when students had a low level of state self-esteem (1 SD below the mean; Aiken & West, 1991), \( b = 0.35, p < .01, p_{\text{rep}} > .95 \), but not when students had a high level of state self-esteem (1 SD above the mean), \( b = 0.05, p > .68, p_{\text{rep}} < .37 \) (see Fig. 1). Thus, we generalized existing laboratory findings to the real world. By contrast, in the self-affirmation condition, narcissism was not associated with increased aggression, regardless of whether students had low or high state self-esteem, \( b = 0.15, p > .22, p_{\text{rep}} < .70 \), and \( b = -0.22, p > .07, p_{\text{rep}} < .86 \), respectively (see Fig. 1). Thus, these analyses indicate that a 15-min self-affirmation writing exercise reduces narcissistic aggression for a period of 1 school week. There was no significant three-way interaction effect involving the long-term intervention dummy variable, \( p < .10, ^*p < .05, ^{*}*p < .01 \).

### Table 1

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Model 1</th>
<th>95% confidence interval</th>
<th>Model 2</th>
<th>95% confidence interval</th>
<th>Model 3</th>
<th>95% confidence interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>0.20**</td>
<td>0.07</td>
<td>0.17*</td>
<td>0.07</td>
<td>0.09†</td>
<td>0.06</td>
</tr>
<tr>
<td>Self-affirmation condition</td>
<td>0.03</td>
<td>0.09</td>
<td>0.03</td>
<td>0.09</td>
<td>0.01</td>
<td>0.07</td>
</tr>
<tr>
<td>Narcissism</td>
<td>0.06</td>
<td>0.04</td>
<td>0.06</td>
<td>0.04</td>
<td>0.02</td>
<td>0.03</td>
</tr>
<tr>
<td>Trait self-esteem</td>
<td>-0.02</td>
<td>0.04</td>
<td>-0.04</td>
<td>0.04</td>
<td>-0.02</td>
<td>0.03</td>
</tr>
<tr>
<td>Female gender</td>
<td>-0.42**</td>
<td>0.09</td>
<td>-0.41**</td>
<td>0.09</td>
<td>-0.19**</td>
<td>0.07</td>
</tr>
<tr>
<td>STI</td>
<td>0.00</td>
<td>0.04</td>
<td>0.03</td>
<td>0.04</td>
<td>0.00</td>
<td>0.05</td>
</tr>
<tr>
<td>LTI</td>
<td>-0.01</td>
<td>0.05</td>
<td>-0.02</td>
<td>0.04</td>
<td>0.00</td>
<td>0.06</td>
</tr>
<tr>
<td>IV</td>
<td>-0.03</td>
<td>0.02</td>
<td>-0.03</td>
<td>0.02</td>
<td>0.00</td>
<td>0.02</td>
</tr>
<tr>
<td>Narcissim × IV</td>
<td>-0.00**</td>
<td>0.02</td>
<td>-0.04†</td>
<td>0.03</td>
<td>-0.05*</td>
<td>0.03</td>
</tr>
<tr>
<td>Narcissim × IV × STI</td>
<td>0.10*</td>
<td>0.04</td>
<td>0.09*</td>
<td>0.04</td>
<td>0.04</td>
<td>0.04</td>
</tr>
<tr>
<td>Narcissim × IV × LTI</td>
<td>0.05</td>
<td>0.05</td>
<td>-0.02</td>
<td>0.04</td>
<td>0.02</td>
<td>0.05</td>
</tr>
</tbody>
</table>

Note. Model 1 predicted aggression from concurrent self-esteem, and Model 2 predicted aggression from self-esteem at the previous assessment; Model 3 predicted self-esteem from aggression at the previous assessment. STI = short-term intervention dummy; LTI = long-term intervention dummy; IV = independent variable. 

1 We also modeled the trajectories of state self-esteem and aggression across the study period by including linear and quadratic time effects. Because these effects were nonsignificant, they were excluded from the subsequent models.

Fig. 1. Results of the analysis testing the intervention effect after one school week. The graph shows the aggression levels of students with low (1 SD below the mean) and high (1 SD above the mean) narcissism and low (1 SD below the mean) and high (1 SD above the mean) concurrent state self-esteem, separately for the self-affirmation and the no-affirmation (control) conditions.
variable, $b = 0.05, p > .31, p_{rep} < .64$; this result indicates that intervention effects dissipated over time. In addition, no effects were found for trait self-esteem. Boys were more aggressive than girls, $b = -0.42, p < .001, p_{rep} > .98$.

In Model 2, we included the same predictors as in Model 1, with the exception that we instead used the state self-esteem level of the previous assessment (i.e., a lagged predictor) as the predictor of aggression (see Table 1). Because we also included an autoregressive term (covariance estimate $= .36, p < .001$, $p_{rep} > .98$), this predictor can be interpreted as the effect of the initial level of state self-esteem on changes in aggression. The results from Model 2 were very similar to the results from Model 1. The two-way interaction between narcissism and state self-esteem did not reach conventional levels of significance, $b = -0.04, p = .09, p_{rep} > .90$. In the control condition, narcissism was associated with increased aggression when students had low but not high levels of lagged state self-esteem, $b = 0.37, p < .01, p_{rep} > .95$, and $b = 0.05, p > .70, p_{rep} < .36$, respectively. In the self-affirmation condition, narcissism was not significantly related to aggression, either for students with lowlagged state self-esteem, $b = 0.09, p > .47, p_{rep} < .53$, or for students with high levels of lagged state self-esteem, $b = -0.24, p > .06, p_{rep} < .87$. Aggression again showed a significant gender effect, $b = -0.41, p < .001, p_{rep} > .98$. In summary, the results from Model 2 extend those from Model 1 by showing that narcissistic aggression follows from initially experienced ego threat, a link that can be temporarily attenuated by means of a self-affirmation exercise. Again, no effects involving the long-term intervention dummy variable or trait self-esteem were found.

To compare the direction of effects for state self-esteem and aggression, we ran an additional model in which state self-esteem was predicted by the aggression level at the previous assessment. In other words, Model 2 showed that state self-esteem was associated with subsequent changes in aggression, and Model 3 tested whether the reverse was also true (Table 1). Not surprisingly, trait self-esteem predicted the average level of state self-esteem (intercept), $b = 0.42, p < .001, p_{rep} > .98$. In addition, a two-way interaction between narcissism and lagged (previous assessment) aggression emerged, $b = -0.05, p < .04, p_{rep} > .89$; narcissistic children tended to react to previous aggression with decreases in self-esteem, perhaps because their self-esteem is sensitive to negative interpersonal encounters (Morf & Rhodewalt, 2001; Thomaes et al., in press). There were no significant effects involving the short-term and long-term intervention dummy variables, a result supporting our prediction that the self-affirmation would influence narcissists’ aggressive responses to lowered self-esteem (rather than narcissists’ self-esteem responses to aggression). A significant gender effect was also found, $b = -0.19, p < .001, p_{rep} > .98$; girls had lower levels of state self-esteem than boys did.

**DISCUSSION**

This randomized field experiment tested whether a self-affirmation intervention can reduce narcissistic aggression in youth. We generalized existing laboratory findings to the real world by showing that narcissistic individuals (not individuals with low self-esteem) behave aggressively when they experience ego threat. More important, we found that this standard pattern was temporarily changed with a short self-affirmation writing exercise. This exercise prevented narcissists from behaving aggressively when they experienced ego threat. Lagged-effects analyses confirmed the predicted direction of effects: The intervention reduced narcissists’ aggression following ego threat (rather than vice versa). The effect of the intervention lasted for a period of 1 school week (i.e., for a period of up to 400 times the duration of the intervention itself).

What accounted for the effectiveness of the seemingly minor self-affirmation in our study? We propose that the self-affirmation temporarily attenuated the ego-protective motivations that normally drive narcissists’ aggression. Previous research has shown that self-affirmations buttress self-esteem. People who are reminded of values that are important to them become less vulnerable to experiences of ego threat, presumably because they realize that their worth as a person does not hinge upon one particular domain of functioning (Creswell et al., 2005; Koole et al., 1999; Sherman & Cohen, 2002). Because vulnerability to ego threat is the key cause of narcissists’ aggressive inclinations (Morf & Rhodewalt, 2001), this previous research suggests that we lessened the motivational source of narcissists’ aggression. Note that the intervention did not raise students’ self-esteem. Trajectories of state self-esteem did not differ following the self-affirmation and control writing assignments. Rather, the intervention made students behave in a less defensive, less aggressive manner when they experienced ego threat.

The developmental timing of the intervention may also have contributed to its effectiveness. Early adolescence is a time when children become increasingly motivated to develop an autonomous identity (Collins & Steinberg, 2006; Erikson, 1968). The intervention allowed the students to reflect on the core values that define them as a person, and so it may have been particularly effective in this developmental period. Finally, small interventions can have strong effects if they interrupt negative cycles of events that would otherwise occur (Cohen et al., 2006). This may well have been the case in our study. Aggressive behaviors rarely are isolated events of one-directional hostility, but often set in motion a sequence of interpersonal hostilities (e.g., Perry, Perry, & Kennedy, 1992; Phelps, 2001). By inhibiting initial outbursts of narcissistic aggression, the intervention may have prevented subgroups of individuals from becoming entrapped in peer conflicts marked by repeated aggressive behaviors.

Our results are consistent with a basic tenet of self-affirmation theory (Sherman & Cohen, 2006; Steele, 1988), namely, that
activities that remind people of “who they are” can have strong behavioral benefits. We have provided the first evidence that those benefits extend to the domain of aggressive behavior. In addition, our results contribute to the debate on the role of low self-esteem that has dominated aggression research for more than a decade (e.g., Baumeister et al., 1996; Donnellan et al., 2005). The weak and inconsistent evidence for a link between low trait self-esteem and aggression has occasionally been attributed to the frequent use of laboratory aggression measures that may not generalize outside the laboratory (Donnellan et al., 2005). Our study examined real-world aggression and still contradicts the view that low trait self-esteem underlies aggression.

The applied relevance of this study is that it provides the first empirical demonstration that buttressing self-esteem can be effective at reducing narcissistic aggression. Two cautions are needed, though. First, the self-affirmation procedure that we used should not be seen as a ready-to-use intervention strategy. We found no evidence for a sustained reduction in aggression longer than a school week. Furthermore, the impact of the intervention was relatively small. It reduced but did not eliminate narcissistic aggression. Future research is needed to generalize our findings to other applied settings, and to explore more powerful self-affirmation intervention procedures that can have longer-lasting effects. Second, the intervention was effective in aggressive youth with narcissistic tendencies, not in aggressive youth in general. Thus, self-affirmation procedures are not likely to be effective as universal, classroom-based aggression interventions.

In conclusion, we hope our study will encourage the development of theory- and evidence-based aggression interventions that target children’s self-views. Many current intervention programs focus on boosting self-esteem, but there are no clear theoretical or empirical reasons why boosting self-esteem should reduce aggression. Self-affirmations buttress self-esteem and buffer people against ego threat, thereby contributing to reducing narcissistic violence and aggression in schools.

REFERENCES


**SUPPORTING INFORMATION**

Additional Supporting Information may be found in the online version of this article:

**Aggression Measure Pilot Study**

**Table S1**

Please note: Wiley-Blackwell is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.