

# No Pain, No Gain: The Conditions Under Which Upward Comparisons Lead to Better Performance

Camille S. Johnson  
Stanford University

Diederik A. Stapel  
Tilburg University

In 3 studies, the authors explored the relation between threatening upward social comparisons and performance. In an initial study, participants were exposed to comparison targets who either threatened or boosted self-evaluations and then completed a performance task. Participants exposed to the threatening target performed better than those in a control group, whereas those exposed to the nonthreatening target performed worse. In Study 2, self-affirmation prior to comparison with threatening targets eliminated performance improvements. In Study 3, performance improvements were found only when the performance domain was different from the domain of success of the comparison target. These boundary conditions suggest that increases in performance following social comparison arise from individuals' motivations to maintain and repair self-evaluations. Implications for the study of the behavioral consequences of social comparison are discussed.

*Keywords:* social comparison, self-affirmation, performance, motivation, self-evaluations

Imagine a typical social comparison situation: John is sitting in a lecture hall waiting to take an exam when he overhears his professor raving about Jane. The professor describes how brilliant Jane is and how wonderful it is to have her in class. There appears to be no end to Jane's accomplishments and achievements. How does hearing about Jane's spectacular attributes make John feel about himself? Does it change his enthusiasm for the course? Does it change his motivation to excel at things he values in life? Will hearing about Jane change John's performance on the exam?

Psychologists have long sought to answer these questions and to understand the many reactions that individuals have to upward

comparisons (Collins, 1996; Wood, 1989). A quick review of the literature suggests, however, that there is no simple answer to whether upward comparisons are good or bad for individuals. Rather, whether an upward comparison benefits individuals depends on the type of response that is measured, the characteristics of the perceiver, and the features of the comparison target. For example, in Tesser's (1988) self-evaluation maintenance model, how social comparisons impact self-evaluations is partly determined by the psychological closeness of target to the perceiver and by the importance of the domain of comparison to the perceiver. Other models, such as Mussweiler's (2003) selective accessibility model, maintain that whether individuals end up feeling similar to or different from the comparison targets depends on whether individuals believe they are similar or different prior to the comparison. Stapel's interpretation comparison model posits that whether upward comparison standards are inspiring or demotivating (lead to assimilation or contrast effects), depends on how perceivers use the comparison information (Stapel & Koomen, 2000). When people use comparison information to define themselves (e.g., to answer the question "Who am I?"), assimilation is more likely, whereas contrast is more likely to occur when the focus is on self-evaluation on a specific dimension (to answer the question, e.g., "How intelligent am I?"; see also Stapel & Suls, 2004).

Given the preponderance of attention dedicated to the relation between social comparison and self-evaluations, our purpose is not to suggest an alternative pathway by which comparisons affect self-evaluations. Rather, our purpose is to build upon this previous research and examine the behavioral consequences of social comparison. Specifically, we seek to further develop previous work examining the relationship between comparison and performance (Dijksterhuis et al., 1998; Mussweiler, 2003; Seta, 1982; Seta, Seta, & Donaldson, 1991). Therefore, we focus on how self-evaluations may contribute to behavior and how self-evaluative and behavioral responses to comparison may be at the same time

---

Camille S. Johnson, Graduate School of Business Behavioral Research, Stanford University; Diederik A. Stapel, Tilburg Institute for Behavioral Economics Research, Tilburg University, Tilburg, the Netherlands.

This research was supported by Grant T32 MH19728 from the National Institute of Mental Health and National Research Service Award F31 MH64238-01 awarded to Camille S. Johnson. Study 3 was completed as part of a dissertation submitted by Camille S. Johnson to the Ohio State University Graduate School in partial fulfillment of the requirements for a doctoral degree. The research and writing of this article was also supported in part by PIONIER grant "Making Sense of Hot Cognition" from the Dutch Science Foundation (*Nederlandse Organisatie voor Wetenschappelijk Onderzoek*) and a research grant of the Heymans Institute of the University of Groningen awarded to Diederik A. Stapel.

We thank Robert Arkin, chair of Camille S. Johnson's dissertation committee, and the other members of the committee, Marilyn Brewer and Richard Petty. We also thank Larissa Tiedens for her comments on an early version of this article and Greg Norman and Valerie Jefferis for their assistance with data collection.

Correspondence concerning this article should be addressed to Camille S. Johnson, Graduate School of Business Behavioral Research, 518 Memorial Way, Stanford University, Stanford, CA 94305, or to Diederik A. Stapel, Tilburg Institute for Behavioral Economics Research, Tilburg University, P.O. Box 90153, 5000 LE Tilburg, the Netherlands. E-mail: johnson.1967@osu.edu or d.a.stapel@uvt.nl

disparate and related. We suggest that the behavioral consequences of upward comparison depend on whether the comparison is threatening to self-regard and whether the performance task provides a means of maintaining self-regard. In short, when upward comparisons threaten self-regard, we predict that improved performance will follow. When upward comparisons do not threaten self-regard, we predict that performance will not be affected. The studies presented here support this argument in two ways.

First, the current article briefly reviews the relevant literature and demonstrates that although there have been some studies that have examined the relation between social comparison and behavior (Dijksterhuis et al., 1998; Mussweiler, 2003; Seta, 1982; Seta et al., 1991), the relations between social comparison, changes in self-evaluations, and performance are less well-understood. The current experiments measure both self-evaluative and behavioral responses to upward comparison side by side and demonstrate how a single comparison target can lead to both lowered self-evaluations and improved performance outcomes (Study 1).

Second, the present experiments outline a number of boundary conditions to the relationship between social comparison and performance that suggest that some upward social comparisons lead to improved performance because they are threatening to self-evaluations. The role of threat in the relationship between comparison and performance is demonstrated using two empirical strategies. First, we demonstrate that when upward comparisons are rendered nonthreatening, performance benefits do not accrue (Study 2). Second, we show that individuals, who are exposed to upward comparison targets that are associated with improvements in performance, act strategically and only increase their striving in performance domains that allow for repair of self-regard (Study 3). Thus, the current studies present a comprehensive picture of how upward social comparisons, threat, and improved performance interact.

## SOCIAL COMPARISON AND BEHAVIOR

Although most social comparison research has focused on the relation between social comparison and self-evaluation, some studies have focused on the impact of social comparison information on behavior. To our knowledge, there have been hardly any studies that consider behavioral and self-evaluative consequences in one empirical design. Moreover, studies of the impact of social comparisons on behavior have not been consistent in their findings. Some studies suggest that upward comparisons should benefit performance, whereas others do not. For example in a classic study, Seta (1982) had participants complete a task with a coactor whose performance was either slightly superior or inferior to their own. Those who worked with the better performing coactor performed better than those exposed to a less successful other. More recently, Marx and Roman (2002) found that women administered a mathematics exam by a female experimenter, who was described as possessing high math ability, performed better than those whose exam was administered by a comparable male experimenter. Additionally, in their investigation of the consequences of selection of comparison targets, Blanton, Buunk, Gibbons, and Kuyper (1999) found that students who selected extremely successful classmates with whom to compare their performance were more successful at the end of the school term than students who compared with

moderately successful classmates. Thus, in these studies the presence of an upward comparison target benefited performance.

On the other hand, other work examining social comparison and performance has found that upward comparison may be harmful for performance. For example, Dijksterhuis et al., (1998) found that when participants were exposed to an extremely intelligent other (Einstein), they performed worse on an intelligence task than when they were exposed to an extremely unintelligent other (Claudia Schiffer). Similarly, when Stapel and Suls (2004) created a situation in which participants compared themselves with an extreme comparison target, they found that implicit comparisons led to contrasts: Upward comparisons led to worse performance, whereas downward comparisons led to better performance.

Thus, the extant work exploring the impact of social comparison on performance has found that upward comparisons can lead to both better and worse performance. Moreover, even among studies reporting congruent positive effects, the mechanisms by which upward comparisons lead to better performance are not entirely consistent. For example, Seta (1982) frames his findings in terms of the resource investment model and argues that superior others increase one's perceptions of the likelihood of success, which subsequently leads to greater effort and better performance. This is similar to Wheeler, Martin, and Suls (1997) proxy model, in which a comparison other may provide information about one's own likely performance. Alternatively, Marx and Roman (2002) describe their work in terms of the presence of a successful other as eliminating stereotype threat. One of the curious qualities of stereotype threat is that awareness of being a member of a negatively stereotyped group may lead individuals to exert more effort, yet lower their performance (Steele, 1997). Thus, the conditions that led to better performance outcomes (the presence of a counter-stereotypical successful other) should also have been the ones that led to relatively less effort and better performance, contrary to the resource investment model. Moreover, in Marx and Roman's studies, the relations among the comparison domain, the target's domain of success, and the perceiver all determined the performance outcome, not just the direction of comparison. Finally, because only members of the negatively stereotyped group benefited from the upward comparison (men were not positively affected by either the male or the female role model), it is not clear whether a more successful other had a boosting effect on performance or removed the performance attenuating effects of a negative stereotype.

The final situation, that described by Blanton et al., (1999), could be construed as showing the impact of upward comparison over time. They asked students to identify with whom they generally compared themselves and used the direction of comparison to predict future outcomes. Unlike other studies in which social comparison and responses are temporally close and in which direction of comparison was assigned, Blanton et al. may have tapped into an ongoing comparison process between target and perceiver in which instrumental information and feedback may have occurred. Thus, the comparisons may have yielded far richer information than the situation of interest to this paper, and to other typical social comparison studies.

Moreover, none of these studies of behavioral social comparison effects reported self-evaluation effects or spoke to the role of self-evaluation. We believe that evaluative and behavioral social comparison effects are intrinsically related and that the failure to

consider the relation between self-evaluations and performance may have obscured the nature of the relation between social comparison and performance. We suggest that changes in self-evaluations lead to changes in performance, and therefore the two must be studied together. Specifically, we argue that when upward comparisons threaten self-regard, individuals may respond to that threat by improving performance. On the basis of this argument, we would attribute the inconsistency in previous research to the absence or presence of threat to self-regard and the absence or presence of opportunities to repair self-regard. Where previous researchers have found that upward comparison increases performance, we suggest that participants both experienced a threat to self-regard and a means for addressing that threat. Where previous researchers have found that upward comparisons harm performance, we suggest that participants either did not experience a threat to self-regard or did not see an opportunity for addressing that threat.

### Threats to Self-Regard

The term *threat* is used by social psychologists to refer to a wide range of psychological experiences. Here, we are primarily concerned with threats to self-regard. Thus, we focus on experiences that may lead individuals to doubt or lower their self-worth. There are a number of ways in which upward comparison targets can pose a threat to self-evaluations. For example, researchers examining assimilation and contrast effects have found that when comparison targets are extreme, individuals tend to contrast their evaluations away from the targets (see Moskowitz & Skurnik, 1999). In this way, extremely superior others lead to lowered self-regard, whereas extremely inferior others lead to improved self-regard. Conversely, individuals tend to assimilate their self-evaluations with moderate targets. In this way, moderately superior others lead to improved self-regard, whereas moderately inferior others lead to lowered self-regard. Thus, extremely superior others could be considered threatening to self-evaluations, whereas moderately superior others would not (Mussweiler, Rüter, & Epstude, 2004; Stapel & Blanton, 2004).

Other research examining perceptions of the comparison targets has identified other qualities of upward comparison targets that yield such targets threatening (Lockwood & Kunda, 1997). In those studies, when individuals felt that they could not attain achievements similar to more successful others, the successful others were not inspirational. Whereas when individuals felt they could attain similar success, more successful others were inspirational. Similarly, Tesser (1988, 2000) found that when individuals felt close to an individual who was successful in a highly valued domain, those individuals were likely to experience threatened self-regard.

Regardless of the means by which an upward comparison threatens self-evaluations, we suggest that those threats to self-evaluations may drive individuals to perform better. When individuals feel threatened by an upward comparison target, they will be motivated to protect or repair their self-evaluations, and one means for self-regard repair is improved performance.

### Repair of Self-Regard

Tesser's self-evaluation maintenance model (1988) describes a number of ways that individuals may deal with threats to self-

evaluations. One strategy is to decrease the importance of the domain in which one has been outperformed. Thus, if an individual is outperformed on a math test, that individual could withdraw from the math domain. The individual could then refocus on performance in an alternative domain, such as literature, in which she could expect to outperform the comparison target. Withdrawing and changing domains following threats to self-regard has also been identified as a means of maintaining positive self-regard by researchers concerned with the negative impact of stereotypes (Crocker & Major, 1989; Steele, 1997).

Consequently, when an individual has just had a threatening social comparison experience, the relation of the performance task to the domain of comparison may determine how much effort one will exert on the task. For example, if an individual has just been outperformed, and is facing the same competitors in the same domain, that task is unlikely to be considered an opportunity for maintaining self-regard, and one would not expect the person to be particularly motivated to work hard on that task. Instead, one might imagine that the individual would want to compete in another domain. Imagine a student, John, has just been outperformed by his girlfriend, Joan, on a chemistry exam. In that situation, John is unlikely to want to compete with Joan in the chemistry domain. Instead, he might prefer an entirely different activity, such as videogames or basketball, or a different intellectual activity that relies upon verbal abilities, such as playing Scrabble. In accordance with this rationale, we suggest that performance on tasks that present opportunities for maintaining self-regard (e.g., those from alternative domains) are more likely to benefit from threatening upward comparisons than tasks that do not provide such opportunities.

## CURRENT STUDIES

The present studies set out to demonstrate that the ways in which social comparisons impact performance are determined by the ways in which those comparisons impact self-evaluations. Specifically, we aim to demonstrate that social comparisons that threaten self-evaluations can also have beneficial effects on performance. Such beneficial effects will ensue when the performance task provides an opportunity to repair self-evaluations. To do so, we present three experiments. Initially, we attempt to reveal the basic relationship between self-evaluation threatening social comparisons and improved performance on a measure of general ability (Study 1). Then, we show that when threat is removed from the comparison situation, performance benefits do not arise (Study 2). Finally, we reveal that it is the opportunity to repair self-regard that governs whether a threatening comparison leads to improved performance. We accomplished this by highlighting the determining role of the relation between the task domain and the comparison target in determining performance outcomes (Study 3).

### Study 1

In Study 1, participants were exposed to a moderate or extreme upward comparison target or to a control condition and completed measures of self-evaluations and performance. The measure of performance was described as a measure of a "new" construct called *integrative orientation intelligence*, which was intended to represent an ambiguously different domain from that of the com-

parison target. In accordance with previous research, extreme targets were expected to lead to lowered self-evaluations, as participants were expected to contrast with the targets; whereas moderate targets were expected to lead to raised self-evaluations, as participants were expected to assimilate to the targets. However, in accordance with our model, in which self-evaluations and performance should be negatively correlated, extreme targets were expected to lead to better performance outcomes while moderate targets were expected to lead to worse performance outcomes.

### Method

#### Participants

Participants were 45 undergraduate students, from all majors, who received partial credit toward a course requirement. Although gender information was not collected, participants were representative of the overall participant pool, which consisted of 70% women.

#### Procedure

Participants were brought to the experimental session to participate in what were described as several unrelated studies. In the first part, participants were asked to read a brief article about a fellow university student who served as the comparison target.

*Comparison targets.* In the moderately successful conditions, the target was described as a moderately successful student (scoring in the top 25% of his classes). In the extremely successful conditions, the target was described as an extremely successful student (scoring in the top 5% of his classes). In both conditions, the comparison target was described as having friends and being socially popular. In the control condition, participants read a brief, neutral article about the university campus. After reading about the targets, participants were asked to take a moment to reflect on the person in the article and try to imagine the person's appearance, habits, and personality.

After reading about the comparison target, participants completed a 3-item self-evaluation measure and a Remote Associates Task (RAT; Mednick, 1962; McFarlin & Blascovich, 1984). Order of the measures was counterbalanced, with half of the participants receiving the self-evaluation measure first and half receiving the performance measure first. The 3-item self-evaluation measure ( $\alpha = .67$ ) asked participants to rate how much they agreed with the following statements on a 5-point scale (1 = *not at all*, 5 = *very much*): "I feel good about myself," "I feel confident about my abilities," and "I feel confident that I understand things."

*Performance measure.* The RAT consists of test items in which participants are given three words (e.g., coffee, cake, butter) and asked to think of a word that relates to the other three (e.g., cup). Participants completed four practice questions (two easy, two difficult) and then 20 test items (10 easy, 10 difficult randomly sorted). The RAT is a relatively ambiguous measure that can be cast as measuring a variety of fictional and real abilities. In this study, the RAT was described as measuring integrative orientation. Participants were informed that "Integrative orientation consists of the ability to see connections between various stimuli and different kinds of information. It is related to one's ability to see solutions

to problems and to solve problems creatively. Individuals who are high in integrative orientation excel at solving difficult dilemmas and considering different types of information at the same time, like doctors, mechanics, technicians, and managers." This fictional ability was chosen because it ambiguously related to the domain of success of the comparison target.

Finally, participants were asked to rate on 5 point scales (1 = *not at all*, 5 = *extremely*) how similar they were to the target, how successful the target was, how intelligent the target was, and how attainable the accomplishments of the target were.

### Results

These studies were concerned with showing that targets that harm self-evaluations can have positive effects on performance. Therefore, participants were presented with threatening and non-threatening comparison targets, and their self-evaluations and performance were measured. Order of presentation of the self-evaluation and performance measures was counterbalanced. Two separate 2 (order of presentation)  $\times$  3 (type of target) analyses of variance (ANOVAs) were conducted to test for order effects. No significant effects of order were found ( $F < 1$ ), therefore we collapsed across order in subsequent analyses.

#### Manipulation Checks

*Extremity.* First, manipulations of extremity were examined. Participants were asked to rate the target's intelligence, success, and the attainability of the target's accomplishment. Results indicated that the extreme comparison target was viewed by participants as more extreme than the moderate target. The extreme target was rated more intelligent ( $M = 4.2$ ,  $SD = .77$ ),  $F(1, 28) = 9.31$ ,  $p = .005$ ; more successful ( $M = 4.73$ ,  $SD = .46$ ),  $F(1, 28) = 17.08$ ,  $p < .001$ ; and less attainable ( $M = 1.6$ ,  $SD = .63$ ),  $F(1, 28) = 141.77$ ,  $p < .001$ ; compared with the intelligence ( $M = 3.47$ ,  $SD = .52$ ), success ( $M = 3.93$ ,  $SD = .60$ ), and attainability ( $M = 4.27$ ,  $SD = .59$ ) ratings of the moderate targets. Additionally, the moderate comparison target ( $M = 4.20$ ,  $SD = .68$ ) was viewed as more similar to the participants than the extreme comparison targets ( $M = 2.20$ ,  $SD = .86$ ),  $F(1, 28) = 512.00$ ,  $p < .001$ .

*Threat.* Having established that the extreme comparison targets were viewed as more extreme, we examined whether extreme comparison targets constituted threatening comparison targets. A one-way ANOVA, examining the participant self-evaluations, revealed a significant effect of type of comparison target,  $F(2, 42) = 21.04$ ,  $p < .001$ . Contrast analyses revealed that participants exposed to the moderate comparison target reported the highest self-evaluations ( $M = 4.38$ ,  $SD = .38$ ) compared with those in the control condition ( $M = 3.62$ ,  $SD = .63$ ),  $t(42) = 4.04$ ,  $p < .001$ , and that extreme comparison targets led to lower self-evaluations ( $M = 3.18$ ,  $SD = .50$ ),  $t(42) = 3.38$ ,  $p = .02$ , compared with those in the control condition. Correlational analysis examined the relationship between participant ratings of the target and their evaluations of themselves and revealed a similar pattern: The more positively they rated the comparison targets (as successful and

intelligent), the less positively they rated themselves,  $r(30) = -.56, p = .001$ .<sup>1</sup>

Because extreme targets were viewed more positively than the moderate targets, we concluded that our manipulation of extremity was effective. The finding that the moderate target positively influenced participant self-evaluations compared with those in the no-comparison control condition and that the extreme target negatively influenced participant self-evaluations led us to conclude that our manipulation of threat was successful. Extreme targets were threatening targets, whereas moderate targets were nonthreatening.

### Main Analysis

We expected that threatening comparison targets would lead to improved performance. A one-way ANOVA revealed that the type of comparison target to which participants were exposed impacted their task performance,  $F(2, 42) = 8.74, p = .001$ . As shown in Figure 1, reading about an extreme comparison target led to performance improvements ( $M = 10.13, SD = 1.46$ ) relative to the control condition ( $M = 9.00, SD = 1.73$ ),  $t(42) = 2.03, p = .049$ . Conversely, reading about a moderate comparison target led to performance decrements ( $M = 7.80, SD = 1.37$ ) relative to the control condition,  $t(42) = 2.15, p = .04$ . Correlational analysis, which examined the relationship between participant ratings of the target and their performance revealed a similar pattern: The more positively they rated the comparison targets, the more items they solved correctly,  $r(30) = .60, p < .001$ . Self-evaluations and performance were negatively correlated,  $r(30) = -.26, p = .08$ .

### Mediational Analysis

A mediational analysis examining the role of self-evaluations in the comparison–performance relationship was conducted. Only participants who engaged in social comparisons ( $N = 30$ ) were included in this analysis. Given the small sample size, we used the bootstrapping method suggested by Preacher and Hayes (2004). The first regression equation revealed that the influence of comparison type (0 = *nonthreatening*, 1 = *threatening*) on performance was significant ( $B = -2.33, SE B = .52$ ),  $t(30) = 4.51, p < .001$ . A second regression equation revealed that self-evaluations significantly predicted performance ( $B = -.68, SE B = .30$ ),  $t(30) = -2.25, p < .001$ . The final regression equation included comparison type in the first step and then added self-evaluations in the second step. When self-evaluations were entered into the equation, the influence of target type on performance remained significant ( $B = 3.52, SE B = .86$ ),  $t(30) = 4.08, p < .001$ . However, a Sobel test (Preacher & Hayes, 2004) revealed that the indirect effect of self-evaluations was marginally significant ( $z = 1.63, p = .10$ ). Thus, the mediational analysis provided some support for our hypothesis that comparisons impact performance because of the effects that comparisons have on self-evaluations.

### Discussion

The results of Study 1 provide an initial demonstration of our predicted effect: Comparisons that threaten self-evaluations boost performance. In response to an extremely successful target, participants lowered their self-evaluations and increased their level of

performance. Moreover, among these participants, their perceptions of the threatening target as more extreme were responsible for their improved performance. Conversely, in response to a moderately successful target, individuals raised their self-evaluations but decreased their level of performance.

Given our overall model that threats to self-evaluations lead to improvements in performance, one might have expected that self-evaluations would have mediated the target–performance relationship yet the indirect effect of self-evaluations was only marginally significant. We do not view this finding as disconfirming our model. The purpose of Study 1 was to demonstrate the disparate effects that can emerge following upward social comparisons: Negative effects on self-evaluations and positive effects on performance. The results, as reported, meet this goal. Moreover, experimental evidence presented in Studies 2 and 3, in which the mediating variable (threat) is manipulated, will demonstrate that it is the threat to self-evaluations caused by certain upward comparison targets that lead to performance increases. As Sigall and Mills (1998) and Spencer, Zanna, and Fong (2005) persuasively argued, in many situations doing experiments may be the preferred method of testing for mediation.

Moreover, this research strategy allows us to retain the organic nature of the social comparison–performance relationship. Although people certainly feel threat and face threatening comparisons each day, rarely are these feelings of threat explicitly expressed or asked about. By demonstrating the role of threat, without inducing our participants to consciously engage in comparison with the targets, we provide a more externally valid and realistic view of how upward social comparisons affect performance.

From the outset, we have argued that threatened individuals would perform better because they perceived the task as a means of self-regard repair. Yet, Study 1 did not find any order effects. That is, participants who viewed an extremely successful comparison target reported equally low self-regard after completing the performance task as they did prior to completing the performance task. There are a number of reasons for this pattern of results. One possible reason is that participants did not receive feedback regarding their performance. Perhaps one must receive positive feedback to experience full repair of self-evaluations. Though they might have anticipated positive feedback, they were not told how they had done on the task. Future studies examining the efficacy of performance as a means of self-regard repair would be useful.

Although the results of Study 1 are consistent with the contention that the experience of threat leads to increased performance, one could argue that differences between self-evaluative responses and performance responses emerge because self-evaluations are subjective measures, whereas performance is an objective measure (Mussweiler, 2003). According to this argument, threats to self-evaluations do not lead to better performance but are simultaneous responses to an extreme comparison target. If this is true, then changes in threats should not impact changes in performance—the two should remain independent. Study 2 tests this hypothesis. Moreover, if differences in responses emerge because one is sub-

<sup>1</sup> Because ratings were administered after the performance measure, a series of ANCOVAs were conducted, controlling for performance. These analyses yielded a similar pattern of results.

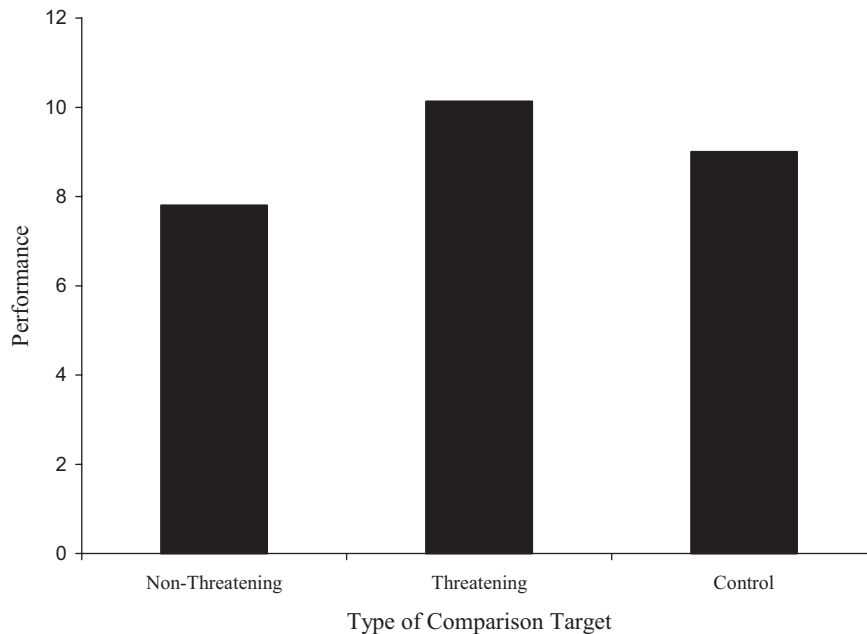


Figure 1. Threatening targets increase performance relative to nonthreatening and control conditions (Study 1).

jective and the other is objective, then changing the domain of the objective measure (performance), should not determine performance. Study 3 tests this hypothesis.

### Study 2

Study 2 demonstrates the role of threat in increased performance following social comparison using a different strategy. First, rather than using extremity to manipulate threat, in Study 2 we used a different means of rendering the comparison target threatening: relative age or attainability. In an often cited study, Lockwood and Kunda (1997) manipulated the age of the comparison targets relative to their participants to create the impression that the accomplishments of older comparison targets were more attainable, and the accomplishments of same-age or younger comparison targets were less attainable. They found that when participants believed the accomplishments of the other to be attainable, participants reported increased self-evaluations. When the accomplishments of the other were unattainable, participant self-evaluations were not boosted by hearing about these accomplishments. Although manipulating extremity in Study 1 had a similar effect (i.e., extreme comparison targets were rated significantly less attainable and led to lowered self-regard, whereas moderate targets were rated more attainable and led to improved self-regard), in theory it is possible that the feeling of threat was not evoked in that study. Therefore, in Study 2, we used the manipulation of threat Lockwood and Kunda used in their studies of the inspiring versus demoralizing effects of role models.

In addition, Study 2 extends Study 1 by manipulating threat to self-evaluations, rather than by simply measuring self-evaluations. In Study 2, participants engaged in self-affirmation prior to social comparison. Self-affirmation has been shown to have buffering qualities, such that engaging in self-affirmation prior to facing a threatening situation can protect self-regard (Steele, 1988; Sher-

man & Cohen, 2002). Consequently, we expect that participants who self-affirm prior to social comparison should not experience threat to self-evaluations and, therefore, should not perform better on a subsequent task. Finally, Study 2 includes measures of effort expenditure. If increased performance occurs because of a desire to repair self-regard, then changes in self-reported effort or task importance should also be found.

### Method

#### Participants

Participants were 84 students in their first 2 years of university. All received partial credit toward a course requirement. Although gender information was not collected, participants were representative of the overall participant pool, which consisted of 70% women.

#### Materials

A 2 (affirmation type: none, affirm)  $\times$  3 (comparison type: none, threatening, or nonthreatening) between-subjects design was used.

In the comparison conditions, participants read about a successful student who had recently won an award. The student was described as Hans de Groot, who had recently won an award at a national student congress. He was described as "one of five finalists chosen . . . to participate in three days of 'intellectual challenges.'" The paragraph also contained quotes from judges of the competition praising Hans' exceptional problem-solving and verbal abilities. Finally, the paragraph gave information about the size of the prize (€4,000). In the threatening conditions, he was described as a first year student and in the nonthreatening conditions he was described as a senior student. It was made clear that Hans

had competed against other students of similar age and standing to win the award.

This manipulation of attainability has been used by other researchers to create threat (Lockwood & Kunda, 1997). In previous research, older targets were associated with increased attainability, increased self-regard and inspiration, while younger targets were associated with decreased attainability (Lockwood & Kunda, 1997). To ensure that our operationalizations were successful, pilot studies were conducted. In our pilot studies, participants rated the older and same-age comparison targets equally creative, successful, likeable, lazy, and extraverted ( $ps > .05$ ). However, older comparison targets were viewed as more attainable, more inspiring, and more similar to the participants than the same-age comparison targets ( $ps < .05$ ). Thus, we judged our manipulation effective and appropriate. Additional analyses of the data collected during the experiment supported this judgment.

### Procedure

Participants were recruited for what they believed to be two separate studies. During the first part of the study, participants were asked to complete a self-affirmation task or no affirmation task at all. In the affirmation condition, they were asked to list four aspects of themselves that show that they are “verbally skilled.” In the no affirmation condition, participants listed four aspects or characteristics of a tree. After completing one of these randomly assigned tasks, participants then read about the comparison targets or read about an unrelated university project.

The targets were described as either an incoming first-year student (threatening) or as a graduating senior (nonthreatening). Participants were asked to take a moment to reflect on the person in the article and try to imagine the person’s appearance, habits, and personality.

Then, participants completed the same performance task and the same self-evaluation measures as in Study 1. The performance task was always completed prior to the self-evaluation task.

Finally, participants were asked to rate on a 5-point scale (1 = *not at all*, 5 = *very much*) how similar to themselves, attainable, successful, and intelligent the comparison targets were. Then, they were asked to rate their own effort on the RAT and how important it was for them to do well on the RAT. These responses were recorded on a 9-point scale (1 = *not at all*, 9 = *very much*).

### Results

We posited that comparison targets that threaten self-evaluations (e.g., unattainable comparison targets) positively affect performance because performance is a means to repair self-evaluations. Thus, processes that prevent self-evaluation threat should also “prevent” increased performance (Tesser, 1988). Therefore, we expected that when participants were given an opportunity to self-affirm prior to reading about threatening comparison targets, no threats to self-evaluations and no improvements in performance should occur.

### Manipulation Checks

To test whether our manipulation of threat was successful, we examined both participant ratings of the comparison target and participant self-evaluations.

*Perceptions of the target.* At the conclusion of the experiment, participants rated how similar to themselves, attainable, successful, and intelligent the comparison targets were. A series of ANOVAs revealed that our manipulations of attainability were successful. As expected, participants perceived the accomplishments of our same-age comparison targets to be less attainable ( $M = 1.67$ ,  $SD = .67$ ) than the older comparison targets ( $M = 4.28$ ,  $SD = .57$ ),  $F(1, 52) = 228.34$ ,  $p < .001$ . Affirmation and the interaction effect of affirmation and target type did not significantly affect perceptions of target attainability ( $F_s < 1$ ). Moreover, affirmation and age of the targets did not affect participants’ perceptions of target success ( $ps > .08$ ) or target intelligence ( $ps > .10$ ). However, the same-age target was seen as less similar ( $M = 2.03$ ,  $SD = .90$ ) than the older target ( $M = 4.04$ ,  $SD = .94$ ),  $F(1, 52) = 59.98$ ,  $p < .001$ . Thus, our manipulation of relative age of the comparison targets successfully manipulated how attainable participants felt that comparison targets’ accomplishments were.<sup>2</sup>

*Threat.* The self-evaluations of participants who did not complete a self-affirmation exercise were also examined. Because self-evaluations were measured after performance, the analysis controlled for actual performance. Analysis comparing both experimental groups to the control condition showed that manipulation of perceived attainability effectively created threat,  $F(2, 80) = 32.75$ ,  $p < .001$ . Further contrast analysis revealed that those exposed to the unattainable comparison target had lower self-evaluations ( $M = 3.06$ ,  $SD = .77$ ) than those in the control condition ( $M = 3.97$ ,  $SD = .66$ ),  $t(81) = 5.45$ ,  $p < .001$ . Additionally, the self-evaluations of those in the control condition were significantly lower than the self-evaluations of those in the attainable condition ( $M = 4.42$ ,  $SD = .44$ ),  $t(81) = 2.58$ ,  $p = .01$ . On the basis of these findings, we judged our manipulation of threat to be successful.

*Self-affirmation.* To confirm the effectiveness of our self-affirmation procedures, we first examined self-evaluations. A 2 (affirmation: none, affirm)  $\times$  3 (comparison type: none, threatening, nonthreatening) ANOVA was conducted. As can be seen in Figure 2, affirmation,  $F(1, 78) = 22.39$ ,  $p < .001$ , and comparison type significantly affected self-evaluations,  $F(2, 78) = 45.92$ ,  $p < .001$ , as did the interaction effect of affirmation and comparison type,  $F(2, 78) = 5.12$ ,  $p = .008$ .

To verify the self-evaluation boosting effect of the self-affirmation exercise, the self-evaluations among participants in the no comparison conditions were examined (the right-most bars in Figure 2). These planned comparisons revealed that the self-affirmation exercise produced the expected boost in self-evaluations. Those who engaged in self-affirmation had higher self-evaluations ( $M = 4.29$ ,  $SD = .55$ ) than those who did not self-affirm ( $M = 3.67$ ,  $SD = .63$ ),  $F(1, 26) = 7.68$ ,  $p = .01$ .

To confirm the self-evaluation buffering effect of the self-affirmation exercise, the self-evaluations of participants in the threatening comparison conditions were examined (the center two bars in Figure 2). Engaging in self-affirmation buffered individuals from the threats self-evaluation posed by threatening comparison targets. When participants self-affirmed prior to seeing the threatening comparison target ( $M = 3.55$ ,  $SD = .50$ ), their self-evaluations

<sup>2</sup> Because ratings were administered after the performance measure, a series of analyses of covariance (ANCOVAs) were conducted, controlling for performance. These analyses yielded a similar pattern of results.

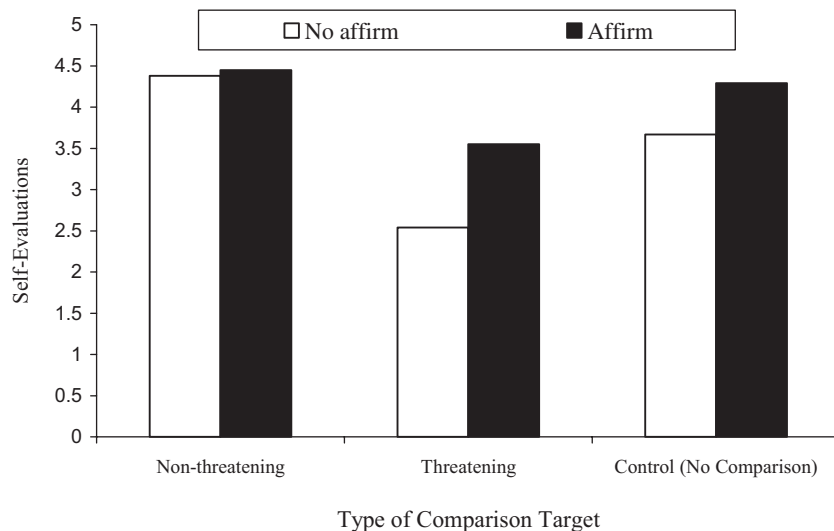


Figure 2. Self-affirmation and nonthreatening upward comparisons have similar effects on self-evaluations (Study 2). Affirm = self-affirmation.

were significantly higher than if they did not affirm ( $M = 2.55$ ,  $SD = .72$ ),  $F(1, 63) = 10.77$ ,  $p = .002$ . Therefore, the affirmation exercise was successful. It raised self-evaluations in the absence of social comparison and buffered participants from self-evaluation threat in the presence of upward threatening comparisons.

### Main Analyses

Threatening comparison targets were expected to lead to improved performance because they threatened self-evaluations.

Therefore, under conditions in which threats to self-evaluations were eliminated (in which participants self-affirmed), social comparison should not have led to improved performance.

We conducted a 2 (affirmation: none, affirm)  $\times$  3 (comparison type: none, threatening, nonthreatening) ANOVA on participant performance to test our hypotheses. As seen in Figure 3, comparison type,  $F(2, 78) = 7.14$ ,  $p = .001$ , and the interaction effect of affirmation and comparison type,  $F(1, 78) = 21.35$ ,  $p < .001$ , significantly affected performance. Further analyses were conducted to test our specific hypotheses.

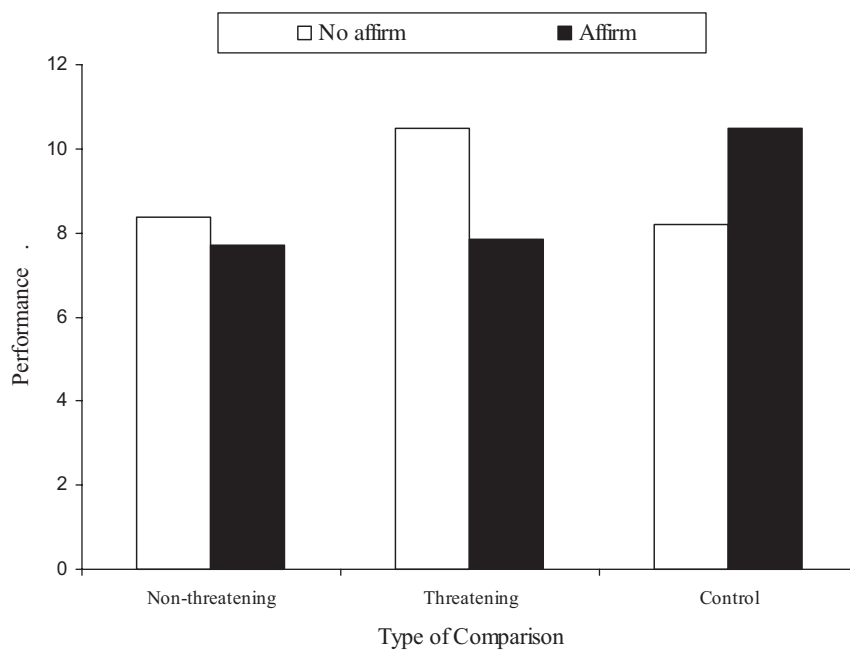


Figure 3. Self-affirmation and threatening upward comparisons have similar effects on performance (Study 2). Affirm = self-affirmation.

In the conditions in which participants did not affirm, the results of Study 1 were replicated. Contrasts revealed that participants who were exposed to the threatening comparison target performed better ( $M = 10.5$ ,  $SD = 8.36$ ) than did participants in either the control condition ( $M = 8.21$ ,  $SD = 1.48$ ),  $t(78) = 4.26$ ,  $p < .001$ , or the nonthreatening comparison target condition ( $M = 8.36$ ,  $SD = 1.34$ ),  $t(78) = 3.99$ ,  $p < .001$ , which did not significantly differ from one another,  $t(78) = -.27$ ,  $p > .05$ .

However, when participants were affirmed, type of comparison target did not impact performance. The performance of those exposed to the threatening comparison target did not differ from that of the participants exposed to the nonthreatening target. However, contrasts revealed that those who were affirmed but did not see a comparison target ( $M = 10.5$ ,  $SD = 1.50$ ) performed better than both the threatening ( $M = 7.86$ ,  $SD = 1.29$ ),  $t(78) = 5.91$ ,  $p < .001$ , and the nonthreatening comparison conditions ( $M = 7.71$ ,  $SD = 1.27$ ),  $t(78) = 4.93$ ,  $p < .001$ . Thus, in the conditions where the threat posed by the threatening comparison target was eliminated, performance improvements were also eliminated.

*Relationship between self-evaluation and performance.* As in Study 1, when people engaged in social comparison, a negative correlation between self-evaluations and performance was expected. However, when participants self-affirmed, boosted self-evaluations were expected to lead to better performance. These effects were found. When participants affirmed (regardless of whether they engaged in comparison), self-evaluations and performance were positively correlated,  $r(42) = .33$ ,  $p = .03$ ; the better participants evaluated themselves, the better they did on the task. When participants saw a threatening comparison (regardless of whether they affirmed), the opposite and expected effect was found: Self-evaluations and performance were negatively correlated,  $r(28) = -.50$ ,  $p = .001$ . The worse participants evaluated themselves, the better they did on the task. Finally, when participants saw a nonthreatening comparison target, performance and self-evaluations were not significantly correlated,  $r(28) = .31$ ,  $ns$ .

*Perceptions of the task.* Participants were asked how important it was for them to do well on the task and how much effort they expended on the task. These responses are reported in Table 1. Two ANCOVAs, controlling for actual performance, were conducted to examine the effects of affirmation and target type on these measures. As expected, those participants whose self-evaluations had been threatened by the threatening comparison target (but had not experienced the buffering effects of self-affirmation) perceived the task to be more important ( $M = 7.71$ ,

$SD = 1.26$ ) than did participants who had been threatened and affirmed ( $M = 5.71$ ,  $SD = 1.27$ ),  $t(78) = 32.44$ ,  $p < .001$ . Moreover, a Tukey's honestly significant difference (HSD) test revealed two homogeneous subsets, such that participants in the threatening target/no affirmation condition viewed the task as more important than all other conditions. Similar interaction effects were found for effort,  $F(2, 77) = 3.57$ ,  $p = .03$ . Participants who were exposed to the threatening comparison target and who did not self-affirm reported expending more effort ( $M = 6.99$ ,  $SD = 1.47$ ) than did similarly threatened participants who did self-affirm ( $M = 5.21$ ,  $SD = 1.81$ ),  $t(78) = 3.26$ ,  $p = .002$ . Again, Tukey's HSD found two homogenous subsets such that threatened and non-affirmed participants differed from all other conditions.

### Discussion

Several results of Study 2 offer insight into the process by which threatening social comparisons positively impact performance. First, the replication of the results of Study 1, with a different manifestation of a threatening comparison target, attests to the robustness of the effect. Less attainable targets led to lowered self-evaluations and better performance, whereas more attainable targets led to raised self-evaluations and worse performance. Second, these effects occurred even though attainability and extremity ratings were not related. The comparison targets that led to better performances in Study 2 were not judged more extreme or more different than the comparison targets that harmed performance. Thus, it appears that it is not merely the extremity of the comparison targets that leads to lowered self-evaluations and improved performance, but it is a larger construct, of which extremity is but one manifestation. In light of the moderating effect of self-affirmation in this study, and the perceptions of the target reported in the pilot study, we suggest that the threat to self-evaluations posed by the target is that larger construct.

In the present study, when self-affirmation was used to eliminate the threat posed by the comparison targets, improved performance was also eliminated. The elimination of the performance benefit that occurred when participants were self-affirmed supports the hypothesis that the experience of threat leads to improved performance. When threat was absent, so were the performance benefits.

Additionally, participants who experienced threat reported that the task was more important and reported expending more effort on that task than did participants who did not experience threat. This suggests that these participants were valuing the task more,

Table 1  
*Participant Perceptions of the Task by Type of Comparison Target (Study 2)*

Perception	Control		Threatening target		Nonthreatening target	
	No SA	SA	No SA	SA	No SA	SA
Effort						
<i>M</i>	5.42 <sub>b</sub>	5.92 <sub>b</sub>	6.99 <sub>a</sub>	5.21 <sub>b</sub>	5.71 <sub>b</sub>	5.67 <sub>b</sub>
<i>SD</i>	1.55	1.33	1.47	1.80	1.56	1.46
Importance						
<i>M</i>	6.36 <sub>b</sub>	6.21 <sub>b</sub>	7.71 <sub>a</sub>	5.71 <sub>b</sub>	6.43 <sub>b</sub>	6.29 <sub>b</sub>
<i>SD</i>	1.15	1.21	1.26	1.27	0.85	0.82

Note. All means not sharing a subscript differ at the  $p < .05$  level. SA = self-affirmation.

and that they viewed integrative orientation ability as important to attaining their intellectual goals. Whether they had experienced threat led them to differentially perceive the task, with those who had been threatened viewing the task in such a way that it would help them protect and maintain their self-regard.

Finally, Study 2 also revealed an interesting pattern of correlations. When participants did not affirm, self-evaluations and performance were negatively correlated: threats to self-regard appeared to be related to better performance. However, when participants did affirm, self-evaluations and performance were positively correlated: boosts to self-regard appeared to be related to better performance. Although the current article is concerned with the consequences of upward comparisons and the threats posed by them, these findings suggest that the means by which self-evaluations are boosted may have different consequences for performance.

Although Study 2 revealed that in the absence of threat, performance improves, one could generate alternative hypotheses by which more successful others might improve performance. For example, previous research has found that the mere presence of others may activate congruent goals (e.g., be neater, work harder) and alter behavior (Fitzsimons, Shah, Chartrand, & Bargh, 2005). Therefore, if social comparison only leads to improvement in the domain in which the comparison target has succeeded, one could argue that performance is improving because the comparison target has activated specific achievement goals. Alternatively, if social comparison led to improvements in performance in all domains, one could argue that comparison or lowered self-evaluations leads to a generalized increase in motivation or a tendency toward action. Furthermore, because the RAT was created as a measure of creativity, and was used in both Study 1 and Study 2, one could argue that the experience of threat is only related to increased performance on creative tasks (Akinola & Mendes, 2007; Friedman & Forster, 2001). Finally, one could still argue that self-evaluative and behavioral responses diverge because self-evaluations are subjective, whereas behaviors are objective measures. Study 3 addresses these issues.

### Study 3

In Study 3, social science/humanities majors were exposed to a successful university student who was either a science or social science/humanities major. Additionally, the success of the science major comparison target was attributed to the target's analytic abilities, whereas the success of the humanities major comparison target was attributed to the target's creative and verbal abilities. For these participants, the humanities major comparison target was considered threatening—the target surpassed the participants in a personally relevant domain. Conversely, the science major comparison target was not expected to be threatening—this target's domain of success was not personally relevant (Tesser, 1988; Johnson & Stapel, 2007). After exposure to the comparison target, participants then completed two tasks: one described as a measure of analytic ability and one described as a measure of creativity.

As in Studies 1 and 2, exposure to a nonthreatening upward comparison target (someone from a different major), was not expected to lead to performance benefits. When participants are exposed to a threatening target, improvements in performance were expected. However, benefits were only expected when the

task provided an opportunity to repair self-regard—when the task was from an alternative domain. When the performance task matched the domain in which participants had just been outperformed, we expected that participants would withdraw from that domain. This hypothesis was derived from Tesser's self-evaluation maintenance model (1988) which suggests that performance in an alternative domain is one means of repairing threatened self-evaluations. Thus, we believed that when the performance domain matched that of the comparison target, no improvements in performance should be expected. Specifically, we expected that performance would be highest when the humanities majors were exposed to a relevant comparison target (a humanities major) and when the task was from an alternative domain (the analytic task).

### Method

#### Participants

Fifty-four (37 male, 17 female) university students participated in this experiment. All participants were recruited for the experiment based on information provided in a mass testing session. First- and second-year students whose declared majors were humanities-related were recruited and received partial credit towards fulfillment of a course requirement. The study was conducted during spring quarter.

#### Materials

We used a 3 (type of target: threatening, nonthreatening, control)  $\times$  2 (type of task: creative or analytic) mixed-factorial design, with type of task as the within-participant variable.

*Comparison targets.* Participants read about a successful second-year student named James Marshall. In all conditions, the target was described as the winner of a prestigious award and \$8,000 prize. Only the domain of success varied by condition.

The science-successful comparison target was described as winning an award "because of his analytical and mathematical abilities" and was praised for "the ease with which [he] organized and integrated scientific information." According to the paragraph, candidates for the award included natural science students, who were described as chemistry, math, biology, engineering, and health sciences majors.

The verbal-successful comparison target was described as winning the award "because of his verbal and creative abilities" and was praised for "the ease with which [he] understood and communicated complicated ideas to others." As with the science-successful target, James was described as socially successful. According to the paragraph, candidates for the award included humanities students, who were described as communications, education, business, and English majors. In all conditions, James was also described as "a considerate and interesting member of the student community." In the control condition, participants read about a university construction project.

In pretests of these materials, participants ( $N = 40$ ) were asked to rate either the science-successful or the verbal-successful target on their logical versus verbal abilities (1 = *higher in verbal abilities*, 9 = *higher in analytical abilities*). They viewed the science-successful targets as being significantly more analytically skilled ( $M = 7.32$ ,  $SD = 2.11$ ) than the verbally-successful targets

( $M = 4.75$ ,  $SD = 2.22$ ),  $t(37) = 3.70$ ,  $p = .001$ . Additionally, participants rated the verbally successful target as more competent on verbal tasks ( $M = 4.60$ ,  $SD = .50$ ) than on logical tasks ( $M = 3.85$ ,  $SD = .67$ ),  $t(18) = 5.40$ ,  $p < .001$ . Participants rated the science-successful target as more competent on logical tasks ( $M = 4.84$ ,  $SD = .50$ ) than on verbal tasks ( $M = 3.63$ ,  $SD = .90$ ),  $t(19) = 3.94$ ,  $p = .001$ .

*Performance measures.* Creative and analytic measures that previously had been validated by other research (Friedman & Forster, 2001) were used here.

*Analytic task.* Participants were told that in future studies, researchers would be interested in how people perform on logic tasks under different questions. Participants were told that they would be completing a brief task providing a measure of their logical reasoning ability. Participants were given six logic questions. These questions were similar to those found in the analytic version of the GRE and were pretested to be solvable by our participant population.

*Creativity task.* Participants were told that in a future study, the researchers would be examining how different environments influence verbal creativity and that the researchers were interested in how creative people were under standard circumstances. Then, participants were instructed to spend 4 min generating creative uses for a brick. Included were instructions to avoid listing common uses and impossible uses. Participants were told that the creativity of their responses would be scored by a panel of experts and that their score would reflect their creative abilities.

Scoring of the responses was conducted in two phases. First, responses collected during pilot testing were ranked from least creative to most creative by trained judges. Then, responses were scored for creativity. Responses that were similar (e.g., "step on to be taller" and "stepstool") were given the same score. Second, new responses obtained in the experiment, but not given previously in the pilot study, were ranked within the context of the responses collected in the pilot study. Thus a single list of all responses ( $N = 459$ ) was created and all responses were rated on a scale of 1 (*not at all creative*) to 10 (*extremely creative*). Then, the average creativity rating of all responses given by a single participant was calculated, yielding a single creativity score. The judges' overall creativity scores were significantly correlated with one another ( $r = .89$ ,  $p < .001$ ).

In pretests of the creativity and logic tasks, university students from all majors ( $N = 40$ ) found the creativity task to be a better measure of creative ability ( $M = 3.64$ ,  $SD = .98$ ) than the logic task, ( $M = 2.41$ ,  $SD = 1.21$ ),  $t(38) = 5.55$ ,  $p < .001$ . Equally, the analytic task was seen as a better measure of logical ability ( $M = 4.33$ ,  $SD = .84$ ) than the creativity task ( $M = 3.49$ ,  $SD = 1.14$ ),  $t(38) = 5.35$ ,  $p < .001$ . These ratings were not affected by whether participants in the pretests read about a successful humanities or science major. However, the ratings of how well each task measured verbal ability were affected by whether the participants read about a science- or verbal-successful target. When participants read about the science-successful target, they rated the creativity task as a better measure of verbal ability ( $M = 2.37$ ,  $SD = 1.30$ ) than when they read about the verbal-successful target ( $M = 1.65$ ,  $SD = .88$ ),  $t(37) = 2.03$ ,  $p = .05$ . Ratings of the logic task were not affected by the type of target.

On the basis of these results, we concluded that the creativity task represented a domain different from the science-successful

comparison target (who was regarded as being more analytically competent) and the logical task represented a domain different from the verbal-successful comparison target (who was regarded as being more verbally competent).

### Procedure

Participants were brought to the laboratory and told that they would be completing two different studies. All materials and measures were presented on computer. The first "study" was described as a study of media influence in which they would be reading a short paragraph and giving their impressions. It was during this phase of the experiment that participants read about the comparison targets.

In the second part of the experiment, participants were told that another researcher was interested in validating tests of logic and creativity. Participants always completed the logic task first, followed by the creativity task. All instructions for each task were presented to participants just prior to beginning the tasks. At the end of the experiment, participants were asked to rate how difficult and how easy the creativity and analytic tasks were on a 7-point Likert-type scale (1 = *not at all*, 7 = *extremely*). Participant ratings of task ease were reverse coded and added to ratings of task difficulty to create index measures of creativity task difficulty,  $r(59) = .67$ , and analytic task difficulty,  $r(59) = .65$ .

### Results

The effect of comparison targets on performance was expected to be determined by the relationship between the domain of success of the comparison target and the domain of the performance task. When the comparison target was from a different major than the participant (and thus nonthreatening), participants were expected to perform equally well on both the analytical and creative task. When the comparison target was from the same major as the participant (and therefore threatening), participants were expected to do better on a task from an alternative domain (analytical) compared with a task from the same domain (creativity). Thus, an interaction effect of type of comparison target and type of test was predicted.

### Main Analyses

To test for this pattern of effects, participant performances on the analytical and creative tasks were subjected to a 3 (type of target: threatening, nonthreatening, control)  $\times$  2 (type of test: analytical, creative) mixed-factorial ANOVA. Prior to analysis, creativity and analytic scores were normalized, and standardized scores were used in the analysis. No significant main effects were found; however, the effect of the interaction was statistically significant,  $F(2, 48) = 3.45$ ,  $p = .04$ . The overall pattern of the interaction can be found in Figure 4. For ease of comprehension, participant performance on the analytic task is depicted in Figure 5, whereas participant performance on the creativity task is depicted in Figure 6.

As expected, participants in the control condition, who did not read about a comparison target, performed equally well on the analytic ( $M = -.29$ ,  $SD = .99$ ) and creativity tasks ( $M = .03$ ,  $SD = .72$ ),  $t(14) = 1.27$ ,  $p > .05$ . Similarly, among participants

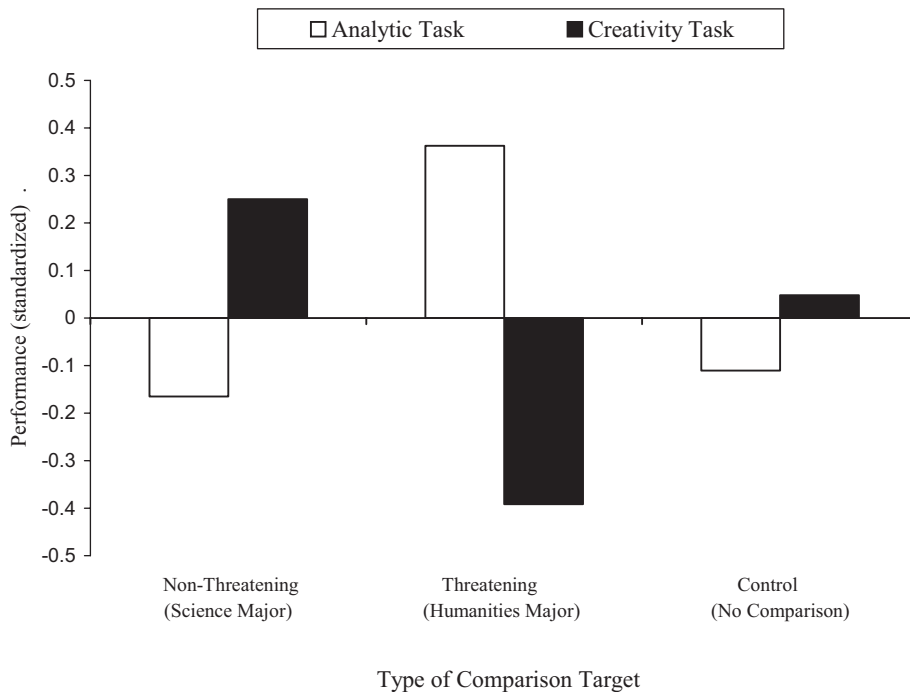


Figure 4. Effect of comparison on performance is determined by the type of task (verbal-creative vs. analytic), standardized scores presented (Study 3).

who read about the nonthreatening science-successful comparison target, performance on the analytic ( $M = -.075, SD = .95$ ) and creativity tasks did not significantly differ ( $M = .22, SD = 1.22$ ),  $t(20) = -.95, p > .35$ .

On the other hand, the performances of those who did read about a threatening comparison differed depending upon the type of task.

These participants performed better on the analytic task ( $M = .39, SD = 1.06$ ) than on the creative task ( $M = -.34, SD = .85$ ). Moreover, contrast analyses revealed that participant scores on the analytic task represented an improvement compared to the control condition ( $M = -.29, SD = .99$ ),  $t(48) = -1.88, p = .06$ . Conversely, on the creativity task participant

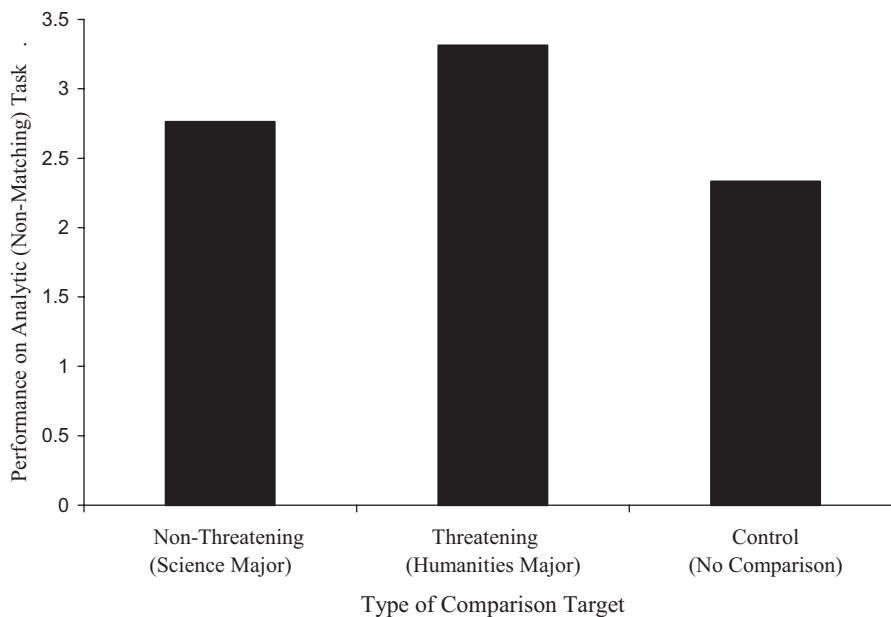


Figure 5. Threatening targets lead to better performance on the mismatching analytic task (Study 3).

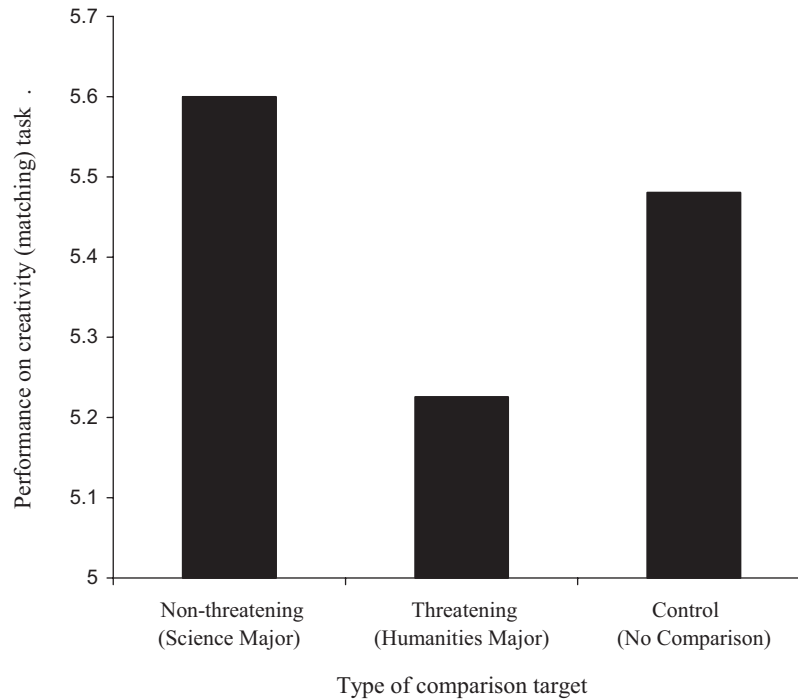


Figure 6. Threatening targets lead to worse performance on the matching verbal creativity task (Study 3).

performance did not differ from the control condition ( $p > .05$ ). Thus, when threatened, participants performed better on a task from a different domain, even when that domain may have been less personally relevant.

#### Perceptions of the Tasks

Participants were also asked about how difficult they perceived the task to be. If comparison targets were serving an instrumental role or as a proxy for the participants (Wheeler et al., 1997), one might expect that science-successful comparison targets would lead participants to view the analytical task as less difficult and that verbal-successful targets would lead participants to view the creativity task as less difficult. This was not the pattern found. Rather, when we controlled for actual performance, participants viewed the creativity task as equally difficult across all three conditions,  $F(2, 47) = 1.54, ns$ . Additionally, when we controlled for actual performance, a marginally significant difference in perceptions of the analytical task was found,  $F(2, 47) = 2.54, p = .09$ . Contrast analysis revealed that participants who saw the verbal-successful comparison target viewed the analytical task as less difficult ( $M = 2.70, SD = .86$ ) than did participants who saw the science-successful comparison target ( $M = 3.31, SD = .78$ ),  $t(56) = 1.71, p = .09$ , or participants in the control condition ( $M = 3.46, SD = .85$ ),  $t(56) = 1.96, p = .06$ . Thus, it appears that comparison targets did not influence perceptions of tasks in a matching fashion. Rather, if anything, exposure to threatening comparison targets (verbal-successful comparison targets) led participants to view tasks from the non-target domains of success as easier.

#### Discussion

The results of Study 3 further illuminate the process by which threatening social comparisons lead to increases in performance. First, participants exposed to threatening comparison targets acted strategically. Had exposure to a more successful person simply activated an overarching goal to succeed, or primed a specific goal to achieve, different results would have been found. In the case of the former hypothesis, task domain would not have mattered, and in the case of the latter hypothesis, better performance on tasks from the same domains would have been expected. Therefore, it appears that participants were performing well in response to threat, not because of increased goal activation or motivation.

Study 3 also refutes the contention that discrepancies in self-evaluative and performance responses emerged in Studies 1 and 2 because self-evaluations are subjective and performance is objective. Here, two objective performance tasks were presented to participants, and it was the relation of the task to the comparison targets domain of success that determined how participants performed. Therefore, it seems less likely that the threats to self and the boosts in performance experienced by participants are two unrelated effects elicited by the characteristics of the measures.

#### GENERAL DISCUSSION

When individuals are faced with upward social comparisons and these comparisons are accompanied by threats to self-regard, they may respond to that threat by increasing their performances. The three studies presented here demonstrate this effect. Using three different types of threatening upward comparisons and two different performance measures, we found that the upward comparison

targets that threatened self-evaluations led to better performances, whereas the upward comparison targets that boosted self-evaluations did not. In addition to providing three demonstrations of the incongruity of self-evaluation and performance effects, the studies also provided evidence that the increase in performance is the result of threats to self-evaluations.

### The Importance of Experiencing and Addressing Threats to Self-Regard

The present studies approached the construct of self-evaluation threat in three different ways. In Study 1, extremity of the target was used as a manipulation of threat because previous research has shown that people tend to contrast with extreme targets. In that study, participants who were exposed to extreme comparison targets lowered their self-evaluations compared with those in a control condition: Extreme upward comparison targets were threatening. In contrast, participants who were exposed to the moderate comparison targets raised their self-evaluations compared with those in a control condition. Although one might have expected that higher self-evaluations would be related to better performance, the participants that experienced self-evaluation threat were the ones that showed improved performance. In that study, the threatening comparison targets were also rated as less similar and less attainable by the participants. Thus, one could argue that the results may have emerged because extreme comparison targets were regarded as outgroup members and less relevant to the participants. However, had extreme targets been completely irrelevant, one would not have expected them to impact participant self-evaluations in a negative manner. Moreover, in Study 2, relative age of the comparison targets, not extremity, was used to manipulate threat, and the same effects were found.

In Study 2, participants saw one of two targets who differed only in their age. Same-age comparison targets led to lowered self-evaluations, whereas older comparison targets led to raised self-evaluations. Thus, same-age targets were threatening while older targets were not. As in Study 1, self-evaluation and performance responses were incongruent. The threatening comparison targets led to better performance outcomes than did the nonthreatening targets. Unlike in Study 1, however, threatening targets were not seen as more extremely intelligent or successful than were the nonthreatening targets. Additionally, threatening targets were rated as more similar to the participants than were the nonthreatening targets. Thus, the two different manifestations of threatening comparison targets in Study 1 and Study 2, drawn from previously published and well-regarded studies, led to the similar effects on self-evaluations and performance but to the different effects on ratings of extremity and similarity. From this, one might conclude that neither extremity nor similarity alone accounts for the performance effects, as they are neither necessary nor sufficient for creating improved performance. The moderating role of self-affirmation further suggests that threat is a necessary component for upward comparisons to lead to improved performance.

In addition to threat, we suggested that an opportunity to repair self-regard was necessary for improved performance to occur. In other words, improved performance occurred by means of a strategic process. Study 3 demonstrated the strategic nature of the performance improvement. In Study 3, yet another instantiation of a threatening target was created. Participants were presented with

either targets who were from the same major or a different major. Participants were also presented with tasks that were either more self-relevant (verbal creativity task) or less self-relevant (analytical logic task). If the comparison target was reminding participants of their own goals and motivations, then one would have predicted better performance on the task that is more congruent with the participants' majors (verbal creativity; Fitzsimons et al., 2005). Equally, if the comparison target was priming a goal to achieve, then one would have predicted better performance on the task that was more congruent with the target's domain of success. Finally, if the comparison target were serving as a proxy, by which participants could judge their own likelihood of success in a particular domain, one would again have predicted better performance on the task that was more congruent with the target's domain of success (Wheeler et al., 1997). However, none of these predictions were borne out. Rather, threatening targets led to better performance on a task that was incongruent with the targets' majors.

One may wonder why a strategic effect occurred in Study 3, but not in Studies 1 and 2. It is important to note that in Studies 1 and 2, the performance tasks were described as measures of "integrative orientation," a globally important ability. Thus, the task was not explicitly linked to the domain of the comparison targets. This ambiguity in the relationship between the comparison domain and the performance domain was purposeful. We allowed participants to view the domain of performance in the most advantageous way possible. Recent research on defensive social comparisons suggests that when under threat, individuals are flexible and strategic in how they view information as relevant to the self (Stapel & Johnson, in press). In those studies, psychology majors were able to take advantage of beneficial comparison information and disregard detrimental comparison information from ambiguously similar others (social science majors). Here, we sought to capitalize on this flexibility by providing participants with performance domains that could be construed as similar or different. In Studies 1 and 2, we created a situation in which it was possible for threatened participants to view the relatively ambiguous performance domain of integrative orientation as different from the domain of success of the comparison target or as a different path to success than that taken by the target.

We argue that it is the difference between the comparison domain and the performance domain that allowed for better performance. One could, however, take the opposite position and argue that because of the ambiguity in the relationship, it was the similarity between the comparison domain and performance domain that allowed for better performance. However, the data from Study 3 supports the former interpretation. In Study 3, we were explicit in defining the performance and comparison and domains as different or the same. In pilot tests of the tasks, the brick task was seen as a better measure of creative-verbal ability than the analytic task, whereas the analytic task was seen as a better measure of logical ability. Moreover, pilot participants exaggerated the differences in the manner that we would expect. Participants who had read about the science-successful target rated the brick task as less related to analytical ability than those who had read about the verbally successful target. Thus, it appears that ambiguity regarding the relationship between comparison domain and performance domain, for the task on which better performance was found, was resolved in such a way that the differences between the domains was augmented and not the similarities.

Although it seems more likely that changing domains should lead to better performance following threatening comparisons than remaining within the same domain, future research should investigate the boundary conditions of this effect. For example, recent research has shown that following the threatening comparison, performance opportunities in different domains can lead to recovery of self-evaluations (Johnson, Norton, Nelson, Stapel, & Chartrand, *in press*) and increases in performance expectations (Johnson & Stapel, 2007b). Future research should examine how similar or different domains must be in order for performance benefits to arise, and investigate how those perceptions of similarity and difference facilitate those performance benefits.

Finally, one may note that in the current studies, we did not explicitly ask participants to compare themselves with the targets. We assumed that both targets were upward comparison targets because they were described as performing well in all of their courses and because participants from this pool have shown themselves to be relatively conservative and realistic when they compare themselves to others (Stapel & Koomen, 2001). However, given self-enhancement biases (Alicke, Klotz, Breitenbecher, Yurak, & Vredenburg, 1995), one could argue that the moderate comparison target in Study 1 was a lateral comparison, not an upward comparison. Although the self-evaluative responses of perceivers to lateral and moderately upward comparison targets are similar, the mechanisms by which these boost self-evaluations may differ. Lateral comparison targets improve self-regard through reassurance (Finlay, Dinos, & Lyons, 2001), whereas moderate targets do so through assimilation (Mussweiler et al., 2004). Thus, one could predict that because different types of comparison targets affect self-regard through different mechanisms, their impact on performance should also differ. The relative effects of lateral and moderate comparison targets could be addressed in future research by explicitly measuring assessments of the comparison targets. Although the focus of the present article is on the discrepancy between self-evaluative and behavioral responses to threatening comparison targets, the results of such investigations could further elucidate the effects of nonthreatening targets.

### Reconciling With Previous Research

Although these findings might at first appear to contradict previous work, because they show that upward comparisons lead to better performance, upon further inspection, one can see that they do not. Instead, this work sheds a different light on previous research. For example, Seta et al. (1991) had participants complete a task adjacent to a coactor, whose performance was either slightly superior or very superior to their own. The task was described to be of either high-value, meaning that a good performance should enhance well-being and status, or of low-value, meaning that good performance should not impact well-being and status. According to the model presented here, an extreme comparison target should have led to better performance on high-value tasks; that is, extreme targets should have threatened self-evaluations, and a high-value task should have provided an opportunity for self-regard repair. However, Seta et al. found the opposite. On high-value tasks, individuals contrasted with the performance of the very superior (extreme) coactor and performed less well than did participants who worked with the slightly superior (moderate) coactors or who worked alone. On low-value tasks, both coactors led to better

performances than working alone. Although our model and Seta et al.'s findings might appear to be contradictory, it is important to note that Seta et al. had participants working on the same task as the comparison target. In their studies, superior targets led to worse performance on important tasks from a matching domain. This is the same pattern observed in the current Study 3. Reframing their findings in terms of our own, it becomes apparent that the present findings and the Seta's resource investment model are complementary and in agreement. This is particularly true if one considers the possibility that being threatened made individuals value performance more (Study 2) and that a shift of performance domain might increase expectations regarding performance, leading to better performance outcomes.

Similarly, Dijksterhuis et al (1998) and Stapel and Suls (2004) found that viewing an extremely superior target led participants to do worse on a test of intelligence. Again, though these effects might, at first, seem to contradict our findings, once one takes into consideration the domain of the comparison target's success and that of the task, the findings become complementary. In those experiments, participants were presented with both an extremely intelligent comparison target and a test of intelligence—the domains were the same. Thus, the likelihood of performing well on the task compared with the target might seem relatively small, leading participants to exert less effort or withdraw from the task. Perhaps had participants been given a different task, such as measuring athletic ability or artistic ability, exposure to an extremely intelligent person might have helped their performances.

This is not to say that our model accounts for all situations in which more successful others boost performance. For example, Marx and Roman (2002) demonstrated that successful female math role models could attenuate stereotype threat effects among female college students and Blanton et al. (1999) demonstrated that choosing more successful others to compare oneself with was associated with better performance outcomes in a classroom situation. In these situations, other influences were present and it is unlikely that threat led to the improved performance outcomes. Indeed, according to Marx and Roman, the presence of the more successful other alleviated stereotype threat. We suggest that the current studies contribute to the understanding of how social comparison and performance are related and suggest that future research should investigate where and how our findings and previous findings intersect. For example, one might argue that threatening targets instigate immediate performance gains through self-defense processes, whereas nonthreatening targets may lead to more beneficial long-term outcomes through changes in goals or aspirations.

### Implications

Implicit in this article is the idea that threats to self-regard are motivating. Certainly, the idea that individuals value self-integrity (Steele, 1988) and positive self-worth (Dunning, 2004) is not new. However, the idea that the protection of self-regard following upward comparisons is tied to specific behavioral outcomes is relatively unexplored. Recently, Lockwood and her colleagues (Lockwood, Kunda, & Jordan, 2002) found that threats to self-regard generated by downward comparisons affected behavioral intentions and levels of motivation. The performance effects revealed here, coupled with Lockwood's work on downward com-

parison, suggest that continued explorations of the relationship between upward comparisons, motivations, and performance would be illuminating.

In the current studies, we suggest that changing domains is required for individuals to perform well following a threatening upward comparison. However, this may be the most extreme situation. That is, individuals may not need to change domains completely but may engage self-defensive mechanisms that allow them to reframe tasks and domains as sufficiently different. For example, a university student might see both verbal ability and logical ability as important for success in her major and might have the overarching goal to be a successful student. When faced with a comparison target whose verbal accomplishments surpass her own, she might begin to see logical ability as more important to success and verbal ability as less important (Beauregard & Dunning, 2001). That is, individuals may not need to switch domains, but may shuffle their perceptions of tasks as more or less related to one another or more or less relevant to different types of goals. This is analogous to social creativity at the group level as discussed in the social identity literature (Jackson, Sullivan, Harnish, & Hodge, 1996; Tajfel & Turner, 1979). When the status of the other person is unattainable or the boundaries of another group are impermeable, individuals may devalue the domain of comparison and view other self- and group-attributes as more valuable. Although not explored in the current studies, these ideas, in addition to the findings of Studies 1 and 2, suggest that individuals need not switch domains entirely in order to experience better performance following threatening upward comparisons. Rather, the data suggest that in order for better performance to result, the task domain must be sufficiently ambiguous such that individuals are able to perceive the performance tasks as distinct from the domain of comparison. That is, individuals may be quite creative in perceiving tasks in advantageous ways that allow for better performance to emerge.

The current findings have a number of implications for research beyond just that which is concerned with threatening comparisons. For example, although the current studies focused on the relation between self-regard threat and performance, Study 2 highlighted an intriguing relation between self-regard boosts and performance. When participants received a boost in self-regard via social comparison, their performance did not improve. However, when participants received a boost in self-regard via self-affirmation, their performance did improve. Thus, two separate processes led to similar effects on self-evaluations, yet led to contradictory effects on performance. A number of reasons for this effect may exist. For example, self-affirmation may boost aspects of self-evaluations related to efficacy and competency, whereas social comparison may boost aspects of self-evaluation related to liking. Future research should examine the importance of the source of self-evaluation changes on performance outcomes (Johnson et al., in press; Johnson & Stapel, 2007a).

Additionally, although we have argued that performance is a means of self-regard repair, we did not show order effects in Study 1. That is, participants who viewed extremely successful comparison targets reported equally low self-regard after completing the performance task as they did prior to completing the performance task. Although this is not entirely surprising given that participants did not receive feedback regarding their performance, future stud-

ies examining the efficacy of performance as a means of self-regard repair would be useful.

## Conclusion

It has been generally assumed that higher self-evaluations are positively associated with higher performance, although recent reviews have suggested otherwise (Baumeister et al., 2003). Equally, it has been postulated that the affective, cognitive, and behavioral responses to a social comparison target should be congruent (Major, Testa, & Bylsma, 1991). This research joins a growing body of work suggesting that these perspectives may be oversimplifications (Stapel & Koomen, 2001; Stapel & Suls, 2004). Rather than being uniform and consistent, the consequences of social comparison may be beneficial or detrimental depending upon the particular response that is measured and the timing of that measurement. These findings suggest that care should be taken when extrapolating from one type of response to another and from one time to another.

In addition, our intention is not to say that popular beliefs about the value of positive role models are misguided. People have long attributed their success to role models who broke barriers and who create new possible selves (e.g., Sally Ride as the first female astronaut), and to role models who provide constant and instrumental feedback (e.g., parents and coaches). Other research on social comparison, particularly the findings of Blanton et al. (1999) and Marx and Roman (2002) may better speak to those specific situations and suggest that when individuals face obstacles or can have meaningful interactions with comparison targets, exposure to more successful others can be beneficial to performance. However, the current findings do speak more broadly to instances in which we seek to influence others and improve their outcomes. They suggest it would behoove us to keep in mind both the wisdom offered by Mark Twain, that "there is nothing so annoying as a good example," and the often touted slogan of physical-fitness fanatics: "No pain, no gain." It appears that though the pain and difficulty of finding oneself less successful than another yields perseverance and improvement, the relative ease and bliss that follows premature feelings of success yield no actual improvements.

## References

- Akinola, M. & Mendes, W. B. (2007). *The dark side of creativity: Affective vulnerability coupled with negative mood leads to greater artistic creativity*. Unpublished manuscript.
- Alicke, M. D., Klotz, M. L., Breitenbecher, D. L., Yurak, T. J., & Vredenburg, D. S. (1995). Personal contact, individuation, and the better-than-average effect. *Journal of Personality and Social Psychology*, *68*, 804–825.
- Baumeister, R. F., Campbell, J. D., Krueger, J. I., & Vohs, K. D. (2003). Does high self-esteem cause better performance, interpersonal success, happiness, or healthier lifestyles? *Psychological Science in the Public Interest*, *4*, 1–44.
- Beauregard, K. S., & Dunning, D. (2001). Defining self-worth: Trait self-esteem moderates the use of self-serving trait definitions in social judgment. *Motivation and Emotion*, *25*, 135–161.
- Blanton, H., Buunk, B. P., Gibbons, F. X., & Kuyper, H. (1999). When better-than-others compare upward: Choice of comparison and comparative evaluation as independent predictors of academic performance. *Journal of Personality and Social Psychology*, *76*, 420–430.

- Collins, R. L. (1996). For better or worse: The impact of upward social comparison on self-evaluations. *Psychological Bulletin*, *119*, 51–69.
- Crocker, J., & Major, B. (1989). Social stigma and self-esteem: The self-protective properties of stigma. *Psychological Review*, *96*, 608–630.
- Dijksterhuis, A., Spears, R., Postmes, T., Stapel, D., Koomen, W., Knippenberg, A. V., & Scheepers, D. (1998). Seeing one thing and doing another: Contrast effects in automatic behavior. *Journal of Personality and Social Psychology*, *75*, 862–871.
- Dunning, D. (2004). *On the motives underlying social cognition*. Malden, MA: Blackwell.
- Finlay, W. M., Dinon, S., & Lyons, E. (2001). Stigma and multiple social comparisons in people with schizophrenia. *European Journal of Social Psychology Special Issue: New Directions in Social Comparison Research*, *31*, 579–592.
- Fitzsimons, G. M., Shah, J., Chartrand, T. L., & Bargh, J. A. (2005). *Goals and labors, friends and neighbors: Self-regulation and interpersonal relationships*. New York: Guilford Press.
- Friedman, R. S., & Forster, J. (2001). The effects of promotion and prevention cues on creativity. *Journal of Personality and Social Psychology*, *81*, 1001–1013.
- Jackson, L. A., Sullivan, L. A., Harnish, R., & Hodge, C. N. (1996). Achieving positive social identity: Social mobility, social creativity, and permeability of group boundaries. *Journal of Personality and Social Psychology*, *70*, 241–254.
- Johnson, C. S., Norton, M. I., Nelson, L., Stapel, D. A., & Chartrand, T. L. (in press). The downside of feeling better: Self-regard repair harms performance. *Self and Identity*.
- Johnson, C. S., & Stapel, D. A. (2007a). *Not all positive self-evaluations are created equal: The consequences of looking inward or outward*. Unpublished manuscript.
- Johnson, C. S., & Stapel, D. A. (2007b). When different is better: Performance following upward comparison. *European Journal of Social Psychology*, *37*, 258–275.
- Lockwood, P. (2002). Could it happen to you? Predicting the impact of downward comparisons on the self. *Journal of Personality and Social Psychology*, *82*, 343–358.
- Lockwood, P., Jordan, C. H., & Kunda, Z. (2002). Motivation by positive or negative role models: Regulatory focus determines who will best inspire us. *Journal of Personality and Social Psychology*, *83*, 854–864.
- Lockwood, P., & Kunda, Z. (1997). Superstars and me: Predicting the impact of role models on the self. *Journal of Personality and Social Psychology*, *73*, 91–103.
- Major, B., Testa, M., & Bylsma, W. H. (1991). Responses to upward and downward social comparisons: The impact of esteem-relevance and perceived control. In J. Suls & T. A. Wills (Eds.), *Social comparison: Contemporary theory and research* (pp. 237–257). Hillsdale, NJ: Erlbaum.
- Marx, D. M., & Roman, J. S. (2002). Female role models: Protecting women's math test performance. *Personality and Social Psychology Bulletin*, *28*, 1183–1193.
- McFarlin, D. B., & Blascovich, J. (1984). On the Remote Associates Test (RAT) as an alternative to illusory performance feedback: A methodological note. *Basic and Applied Social Psychology*, *5*, 223–229.
- Mednick, S. (1962). The associative basis of the creative process. *Psychological Review*, *69*, 220–232.
- Moskowitz, G. B., & Skurnik, I. W. (1999). Contrast effects as determined by type of prime: Trait versus exemplar primes initiate processing strategies that differ in how accessible constructs are used. *Journal of Personality and Social Psychology*, *76*, 911–927.
- Mussweiler, T. (2003). Comparison processes in social judgment: Mechanisms and consequences. *Psychological Review*, *110*, 472–489.
- Mussweiler, T., Rüter, K., & Epstude, K. (2004). The ups and downs of social comparison: Mechanisms of assimilation and contrast. *Journal of Personality and Social Psychology*, *87*, 832–844.
- Preacher, K. J., & Hayes, A. F. (2004). SPSS and SAS procedures for estimating indirect effects in simple mediation models. *Behavior Research Methods, Instruments & Computers Special Issue: Web-based archive of norms, stimuli, and data: Part 2*, *36*, 717–731.
- Seta, J. J. (1982). The impact of comparison processes on coactors' task performance. *Journal of Personality and Social Psychology*, *42*, 281–291.
- Seta, J. J., Seta, C. E., & Donaldson, S. (1991). The impact of comparison processes on coactors' frustration and willingness to expend effort. *Personality and Social Psychology Bulletin*, *17*, 560–568.
- Sherman, D. K., & Cohen, G. L. (2002). Accepting threatening information: Self-affirmation and the reduction of defensive biases. *Current Directions in Psychological Science*, *11*, 119–123.
- Sigall, H., & Mills, J. (1998). Measures of independent variables and mediators are useful in social psychology experiments: But are they necessary? *Personality and Social Psychology Review*, *2*, 218–226.
- Spencer, S. J., Zanna, M. P., & Fong, G. T. (2005). Establishing a causal chain: Why experiments are often more effective than mediational analyses in examining psychological processes. *Journal of Personality and Social Psychology*, *89*, 845–851.
- Stapel, D. A., & Johnson, C. S. (in press). When nothing compares to me: How defensive motivations and similarity shape social comparison effects. *European Journal of Social Psychology*.
- Stapel, D. A., & Koomen, W. (2000). Distinctness of others and malleability of selves: Their impact on social comparison effects. *Journal of Personality and Social Psychology*, *79*, 1068–1087.
- Stapel, D. A., & Koomen, W. (2001). I, we, and the effects of others on me: How self-construal level moderates social comparison effects. *Journal of Personality and Social Psychology*, *80*, 766–781.
- Stapel, D. A., & Suls, J. (2004). Method matters: Effects of explicit versus implicit social comparisons on activation, behavior, and self-views. *Journal of Personality and Social Psychology*, *87*, 860–875.
- Steele, C. M. (1988). The psychology of self-affirmation: Sustaining the integrity of the self. In L. E. Berkowitz (Ed.), *Advances in experimental social psychology*, Vol. 21: *Social psychological studies of the self: Perspectives and programs* (pp. 261–302). San Diego, CA: Academic Press.
- Steele, C. M. (1997). A threat in the air: How stereotypes shape intellectual identity and performance. *American Psychologist*, *52*, 613–629.
- Tajfel, H., & Turner, J. C. (1979). An integrative theory of intergroup conflict. In W. G. Austin & S. Worchel (Eds.), *The social psychology of intergroup relations* (pp. 33–47). Monterey, CA: Brooks/Cole.
- Tesser, A. (1988). Toward a self-evaluation maintenance model of social behavior. In M. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 21, pp. 181–227). New York: Academic Press.
- Tesser, A. (2000). On the confluence of self-esteem maintenance mechanisms. *Personality and Social Psychology Review*, *4*, 290–299.
- Wheeler, L., Martin, R., & Suls, J. (1997). The proxy social comparison model for self-assessment of ability. *Personality and Social Psychology Review*, *1*, 54–61.
- Wood, J. V. (1989). Theory and research concerning social comparison of personal attributes. *Journal of Personality and Social Psychology*, *106*, 231–248.

Received January 19, 2006

Revision received November 12, 2006

Accepted November 12, 2006 ■