European Journal of Social Psychology
Eur. J. Soc. Psychol. **37**, 258–275 (2007)
Published online 18 August 2006 in Wiley InterScience (www.interscience.wiley.com). **DOI**: 10.1002/ejsp.361

When different is better: Performance following upward comparison

CAMILLE S. JOHNSON[1]* AND
DIEDERIK A. STAPEL[2]
[1]*Stanford University, USA*
[2]*University of Groningen, The Netherlands*

Abstract

The attainability of upward social comparisons is known to affect self-evaluative responses. The consequences for performance*, however, are less well understood. We suggest that demoralizing upward comparisons with unattainable targets may lead to improved performance when the target and performance domains are mismatched. For example, comparison with a target that has been successful in an analytic domain should lead to better performance in a verbal domain. This improvement in performance occurs because increased performance in alternative domains provides an opportunity for self-evaluation maintenance. In three studies, we demonstrate that upward comparisons to targets whose successes are perceived as threatening lead to improved performance when the task and performance domain do not match, but no improvements when the domains match. Copyright © 2006 John Wiley & Sons, Ltd.*

In judging ourselves, we look to the accomplishments of others. Their achievements serve as comparison points by which we measure our own progress (Festinger, 1954). In addition to providing information, comparison with another individual may affect our feelings, thoughts, and behaviors. Social comparison researchers have effectively documented how others affect our self-evaluations. Researchers have focused on the direction of comparison (Wood, 1989), the type of comparison (Stapel & Suls, 2004), the role of individual differences (Marx, Stapel, & Muller, 2005), the characteristics of the comparison target (Lockwood & Kunda, 1997) and the conditions under which comparison occurs (Mussweiler, Ruter, & Epstude, 2004). Additionally, researchers interested in performance have shed light on the relation between social comparisons and performance (Blanton, Buunk, Gibbons, & Kuyper, 1999; Marx & Roman, 2002; Seta, 1982; Seta, Seta, & Donaldson, 1991). However, the

*Correspondence to: C. S. Johnson, Graduate School of Business, Stanford University, 518 Memorial Way, Stanford, CA 94305, USA. E-mail: camille_johnson@stanford.edu

Contract/grant sponsor: National Institute of Mental Health, and National Research Service Award F31 MH64238-01; contract/grant number: T32 MH19728.
Contract/grant sponsor: Dutch Science Foundation (Nederlandse Organisatie voor Wetenschappelijk Onderzoek).
Contract/grant sponsor: Heymans Institute of the University of Groningen.

Received 5 April 2005
Accepted 24 March 2006

underlying relation among social comparison, self-evaluations and performance is less well explored. In this paper, we investigate how self-evaluative responses to upward comparison targets lead to strategic performance improvements.

We begin with a distinction first offered by Lockwood and Kunda (1997): the perceived attainability of comparison targets. Lockwood and Kunda exposed participants to targets who were either younger or older than the participants. Older comparison targets represented attainable comparison, because participants could foresee similar accomplishments, whereas younger comparison targets were considered unattainable. They found that attainable comparison targets inspired participants and raised self-evaluations, whereas unattainable targets demoralized participants and lowered self-evaluations. We extend these findings to performance outcomes and make what, at first, may appear to be an ironic prediction. We suggest that under some conditions, the same comparison targets that lead to lowered self-evaluations may lead to improved performance and that the targets that boost self-evaluations may fail to impact performance at all.

This prediction is based on the reasoning that although individuals may not have an opportunity to surpass the accomplishments of an unattainable target in one domain, the possibility of surpassing the target in *other* domains exists. Moreover, the unattainable comparison target threatens self-evaluation and that threat must be resolved (Tesser, 1988). One means of responding to this threat is to increase performance in alternative or mismatching domains. Thus, we predict unattainable comparison targets will lead to improved performance, but only in domains that *mismatch* the target's domain. However, attainable targets are not threatening to self-evaluations. In fact, they may even be inspiring and ego-boosting (Lockwood & Kunda, 1997). Thus, following comparison with attainable targets, there is no threat to be resolved, and the domain of performance should not be relevant: no performance improvement is expected, regardless of the domain.

Several social psychological theories support our hypothesis. According to the self-evaluation maintenance (SEM) model, being outperformed by another may threaten self-evaluations (Tesser, 1988). For example, if a first-year chemistry student compares herself with an outstanding fellow first-year chemistry student, her self-evaluations may be threatened. In the face of this threat, the SEM suggests that she may use one of several strategies to maintain her self-evaluations. One means of dealing with this threat to self-evaluations would be to improve her chemistry performance. However, having been outperformed once in the chemistry domain, rewarding future comparisons in that domain may not be likely and further comparisons in the same domain may only confirm incompetence. On the other hand, comparison with the target in an alternative domain, such as artistic expression, may present an opportunity for superior performance and restoration of self-evaluations. Withdrawing from domains in which further threats to self-evaluations may be anticipated and seeking alternative domains in which comparisons may be positive allows one to maintain positive self-regard (Crocker & Major, 1989; Steele, 1997). If our chemistry student leaves the chemistry domain, she is less likely to face more unflattering upward comparisons and more likely to encounter self-enhancing downward comparisons. Applying this logic to the experimental situation, we expect that our participants, when presented with a task from a domain other than the target's domain of success, will show improved performance. However, when the performance domain matches the target's domain of success, no change in performance is expected.

As we move through daily life we encounter successful others from many different domains, which may match, mismatch, or encompass our own areas of interest. For example, our chemistry student is as likely to encounter a successful chemistry student as she is to encounter someone who is globally or generally successful. The globally successful student may constitute an ambiguous match—she neither matches nor mismatches our chemistry student's domain. In fact, the domain of the globally successful student may be described as encompassing the chemistry domain. What are the consequences of comparison with such a global target? We suggest that in this situation, our chemistry

student can adopt 'socially creative' strategies in which she could carve out a specific domain in which she could be successful. This suggests that when a target's success is not restricted to a particular domain but is described as global or ambiguously matching, individuals may view any specific domain as an opportunity to improve, and performance in that specific domain will improve following comparison. That is, after seeing a globally successful student, individuals are expected to perform better on verbal, analytical, or other specifically described task.

CURRENT STUDIES

The current studies seek to demonstrate that individuals respond to unattainable comparison targets as they would to other self-evaluation threats. In particular, we predict that following exposure to an unattainable comparison target, individuals will increase performance in domains that mismatch the domain of the comparison target. By contrast, attainable comparison targets are expected to boost immediate self-evaluations, but not affect performance.

Three studies are presented here. Building on a foundation provided by previous research (Lockwood & Kunda, 1997), the current studies use manipulations of attainability to create threats to self-evaluations. However, in addition to examining how attainability impacts self-evaluations, the current studies also examine how attainable and unattainable comparison targets impact performance. Thus, rather than extrapolating from self-evaluative responses to predict behavior, the current studies provide evidence of the actual performance consequences. In all studies, participants read about comparison targets who were either the same age (unattainable) or older (attainable), then completed measures of performance or self-evaluation. Each study provides support for a different aspect of our model.

The first study examined the effects of strict matching (i.e., target excels in literature and verbal ability is tested) and mismatching (i.e., target excels in mathematics and verbal ability is tested) of domains on performance. The second study directly examines the relationship among social comparisons, test domains, performance expectations and performance. In addition to measures of performance, Study 2 includes measures of performance expectations and tests the hypothesis that alternative domains lead to better performances when individuals have been exposed to unattainable comparison targets, because alternative domains allow individuals to have higher performance expectancies. The third study examines the generalizability of these effects by examining the effects of matching (i.e., target excels in literature, verbal ability is tested) versus overlapping of domains (i.e., target is globally successful, verbal ability is tested) on performance. Study 3 tests the hypothesis that individuals who experience threat will respond to an encompassing mismatching condition as mismatching, and when exposed to an unattainable comparison target, will experience an improvement in performance.

STUDY 1

Study 1 varies the degree to which the target and test domain match. We suggest that better performance is the result of self-evaluation maintenance behaviors and will only occur when both threats to self-evaluations and opportunities to recover self-evaluations are present. Therefore, better performance is predicted when the participants are exposed to a threatening, unattainable comparison target and provided with an opportunity to perform in an alternative domain.

In addition to performance, participant perceptions of the test are also measured. When participants are doubtful or concerned about their ability level in the test domain, they are expected to view the performance task as more difficult. Thus, we predict that participants who have been exposed to the unattainable comparison target and are faced with a task in a matching domain will anticipate that test as being difficult. After taking the test, those who used the testing as an opportunity to repair self-regard (unattainable/mismatch condition) should also feel greater satisfaction with their performance.

PILOT STUDY

Although previous published research has found that older comparison targets are regarded as attainable and inspiring and that younger comparison targets are regarded as unattainable and discouraging (Lockwood & Kunda, 1997), given the importance of that construct in our theory, it was necessary to verify that the comparison targets used in our studies similarly affected feelings of inspiration and threat. Moreover, given our hypotheses regarding domain-specific influences of older and younger comparison targets, it was necessary to demonstrate that disparate effects of older and younger comparison targets on performance expectations. Therefore, a pilot study of our experimental materials was conducted.

Forty-two Dutch university students read about a successful fellow student. All materials were presented in Dutch, and translations are presented here. The paragraph outlined the process by which Hans de Groot won a (fictional) prestigious prize. It described Hans as 'one of five finalists chosen . . . to participate in three days of ''intellectual challenges'' '. The paragraph also explained that Hans won the prize because he showed 'remarkable intellectual ability during the completion of a variety of tasks' and 'especially because of his verbal ability'. The paragraph also contained quotes from judges of the competition praising Hans' exceptional talent for seeing 'linguistic, verbally well thought out solutions in everyday problems'. Finally, the paragraph gave information about the size of the prize (4000 euros). In the older comparison target conditions, the paragraph ended by reiterating the fact that Hans competed against other advanced students, and in the younger target conditions, the paragraph reminded participants that Hans competed against beginning university students.

After reading about the comparison target, participants were asked to rate the target along a series of traits, including likeability, intelligence, success, creativeness, outgoing, and laziness. Age of the comparison targets did not affect any of these ratings, $t < 1.0$. Successful younger targets were not rated as more successful or more liked than older targets.

Participants also rated how similar the comparison targets were to themselves, how attainable they were, how threatening, and inspiring. Again, the expected effects were found. Participants found the older targets to be more attainable ($M_1 = 4.38$, $t(41) = 2.43$, $p = 0.02$), inspiring ($M_2 = 3.33$, $t(41) = 2.55$, $p = 0.01$), and similar ($M_3 = 4.90$, $t(41) = 2.12$, $p = 0.01$) than the younger targets ($M_1 = 3.59$, $M_2 = 2.45$, $M_3 = 4.05$). Additionally, older targets were viewed as less threatening ($M = 5.77$) than the younger comparison targets, ($M = 5.00$), $t(41) = -2.97$, $p = 0.005$.

Participants were also asked two questions regarding their expectations of future success: to what extent they expected to be successful in their own way and to what extent they expected that they could be successful in the same way as the target. As predicted, participant ratings of the likelihood of success in their own way did not differ as a result of exposure to an younger role model ($M = 4.68$) or to an older target ($M = 4.62$), $t < 1.0$. However, participant ratings of the likelihood of success in the target's domain was affected by the age of the target. When reading about a younger comparison target, ratings of success in the target's domain were statistically significantly lower ($M = 4.05$) than when reading about an older comparison target ($M = 4.86$), $t(41) = 2.42$, $p = 0.02$.

The results of the pilot study suggested that our manipulation of threat—comparison with older or younger comparison targets, is appropriate. Moreover, they provide evidence that people feel that it is less likely that they will succeed in the same domain as a successful other, perhaps leading to a shift in domains.

Method

Participants

Participants were 98 Dutch university students and all materials were presented in Dutch. All students 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% females. All participants were undergraduate students in their first two years of school.

Design

A 2 (type of comparison target: attainable versus unattainable) $\times$ 2(type of domain: domain match versus domain mismatch) between-participants design was employed. Participants viewed the comparison targets then completed the performance task.

Materials

Comparison Target

The comparison targets were the same as used in the pilot study except that the descriptions varied in the abilities to which the targets' successes were attributed. In the matching conditions, the target's success was ascribed to his verbal abilities (literature domain) and in the mismatching conditions to his analytical abilities (mathematics domain). In mismatching conditions, where participants read about the target who was successful because of his mathematics ability, participants were told that Hans had been successful 'especially because of his analytical ability'. The paragraph also contained quotes from judges of the competition praising Hans' exceptional talent for seeing 'logical, analytically well-thought out solutions in everyday problems.'

Performance Measure

The test was described to participants as a measure of verbal ability. The test consisted of 20 remote associate task items (RAT, McFarlin & Blascovich, 1984). Typical RAT items present three related words (e.g., *television*, *window*, and *computer*) and ask participants to fill in a fourth related word (e.g., *screen*).

Procedure

Participants were brought to the lab to participate in several unrelated studies. They were asked to read about the comparison target as part of a study of media influence. After reading about the comparison target, they were asked to take a moment to reflect on the target.

After reading about the comparison target, participants received test instructions. In all conditions the performance task was described as a measure of verbal ability. Participants were told that they should work on the test until they either answered all of the questions or felt that they could not answer any more questions correctly.

After seeing the test instructions, participants rated how difficult the test appeared, then completed the RAT. After the test, participants were asked how satisfied they were with their performance and how difficult the test was for them. All ratings were recorded on 7-point scales (1 = not at all and 7 = very much).

RESULTS AND DISCUSSION

Performance

When the comparison target was unattainable and when the test was in a domain that mismatched the target's expertise, participants were expected to perform better. Therefore, participants in the unattainable and mismatch condition were expected to solve more problems correctly than participants either in the matching or in the attainable conditions. A marginally statistically significant main effect of target on performance was found, $F(1, 94) = 3.5$, $p = 0.07$. Exposure to unattainable targets ($M = 7.65$) led to better performance than exposure to attainable targets, ($M = 6.98$). Additionally, mismatching ($M = 7.80$) led to better performance than matching ($M = 6.78$), $F(1, 94) = 8.47$, $p = 0.005$. These main effects were qualified by the predicted statistically significant interaction, $F(1, 94) = 6.60$, $p = 0.01$. When the target and test domain mismatched, unattainable targets ($M = 8.62$) led to better performance than attainable targets ($M = 7.04$), $F(1, 97) = 8.93$, $p = 0.004$. When the test domain matched, unattainable ($M = 6.67$) and attainable targets ($M = 6.92$) led to similar levels of performance, $F < 1$. Contrast analysis revealed that participants in the unattainable mismatch condition performed better than in all other conditions, $F(1,94) = 13.6$, $p < 0.001$. (See Figure 1 and Table 1).

Perceptions of Test

In addition to affecting performance, target attainability and domain matching were expected to influence perceptions of the testing situation. As with performance, domain matching was only expected to be influential when participants were exposed to the unattainable target. All means are presented in Table 1.

A main effect of type of target for anticipated test difficulty was found ($F(1,94) = 38.41$, $p < 0.001$). Attainable targets led to perceptions of lower test difficulty ($M = 4.25$) compared to unattainable targets ($M = 3.10$). This main effect was moderated by a statistically significant interaction effect, $F(1, 94) = 5.32$, $p = 0.02$. As expected, participants in the unattainable condition anticipated that the test that matched the target domain of success would be more difficult ($M = 4.58$) than the test that mismatched the target domain ($M = 3.92$), $F(1,96) = 4.55$, $p = 0.04$. However, among participants in the attainable condition, test difficulty did not differ, $F < 1$.

Ratings of difficulty and satisfaction following the test were also examined, controlling for actual performance. Again, a main effect for type of target emerged ($F(1,94) = 42.17$, $p < 0.001$), which was moderated by the interaction effect, $F(1,94) = 5.46$, $p = 0.02$. Among participants exposed to the unattainable target, those in the match condition rated the test as more difficult ($M = 4.54$) than those in the mismatch condition, ($M = 3.75$). Again, among participants in the attainable condition, ratings did not differ, $F < 1$.

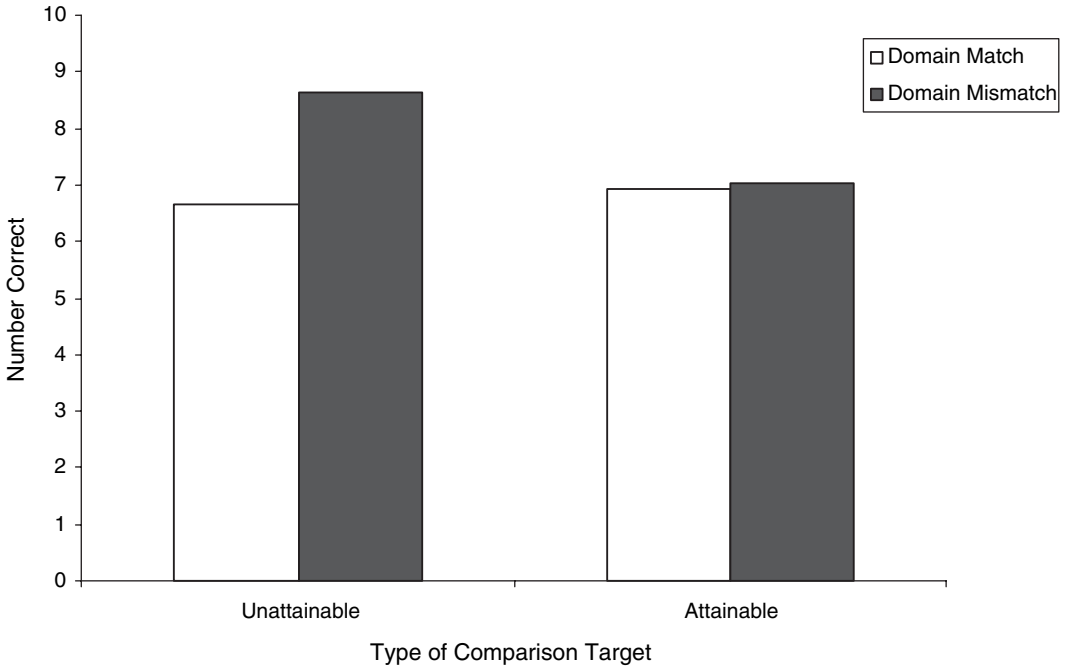

Figure 1. Number correct by role model attainability and matching of domain (Study 1)

Participant ratings of satisfaction were expected to mirror ratings of difficulty. A main effect of domain matching was found ($F(1,94) = 22.22$, $p < 0.001$), which was moderated by the expected interaction effect, $F(1,93) = 26.75$, $p < 0.001$. Among participants exposed to the unattainable target, mismatching led to greater satisfaction ($M = 4.75$) than matching of domains ($M = 2.67$), $F(1, 94) = 39.09$, $p < 0.001$. Again, in the attainable condition, no differences were found, $F < 1$.

Finally, the relationship between test perceptions and actual performance were examined. Prior expectations of test difficulty, post-test ratings of test difficulty, and satisfaction were unrelated to performance on the test.

As expected, when the target and test domains matched, the test was rated as more difficult both before and after completion. Additionally, matching led to greater satisfaction with the test situation. These ratings may reflect a more positive view of the test situation.

Table 1. Means (and standard deviations) of participant perceptions of the task by comparison target and matching of domain (Study 1)

	Unattainable target			Attainable target		
	Match	Mismatch	Overall	Match	Mismatch	Overall
Performance	6.67^{b} (1.63)	8.62^{a} (1.38)	7.65 (1.79)	6.92^{b} (1.79)	7.04^{b} (2.14)	6.98 (1.96)
Prior difficulty	4.54^{a} (0.88)	3.92^{b} (1.28)	3.1 (0.71)	3.00^{c} (0.66)	3.19^{c} (0.75)	4.25 (1.14)
Post difficulty	4.54^{a} (1.02)	3.84^{b} (0.76)	4.58 (1.02)	2.83^{c} (0.76)	3.04^{c} (0.77)	3.75 (1.15)
Satisfaction	2.67^{a} (1.34)	4.75^{b} (1.23)	3.71 (1.65)	3.79^{c} (1.06)	3.65^{c} (1.02)	3.72 (1.03)

Note: Means not sharing a superscript differ at the 0.05 level or less.

Although our hypotheses regarding performance and perceptions of the test were supported, this study did not directly measure self-evaluations. Therefore, Study 2 includes both measures of self-evaluations and measures of performance and directly measures the mediational relationships predicted by our model.

STUDY 2

Study 1 demonstrates that comparison targets that are inspiring are not necessarily good for performance and that comparison targets that are demoralizing are not necessarily bad for performance. Study 2 examines more closely the mechanism by which this transformation from threat to productivity is wrought. To do so, additional dependent variables, as well as an additional control condition are included. First, Study 2 measures participants' self-evaluations directly. Second, the study includes measures of performance threat and performance expectations. By doing so, we can test the relationship between shifting domains and rising performance expectancies. Specifically, a moderated mediation analysis is used to test the hypothesis the way in which comparison target affects performance is moderated by the type of test, and that the change in participant performance expectations is responsible for that moderation (Muller, Judd, & Yzerbyt, 2005). That is, when participants are exposed to an unattainable comparison target, and are given the opportunity to perform on a mismatching task, their expectations regarding the task should lead to better performance.

Study 2 extends the findings of Study 1 in one other important way. In Study 1, the match between target and test domain was manipulated by changing attributions for the target's success. This change could also have affected identification with the target or with the target's domain. Although results of the pilot study suggest that this alternative hypothesis is unlikely, to avoid this possibility, we decided to manipulate the description of the test in order to manipulate the match between target and test domain.

Method

Participants

Participants were 91 Dutch university students and all materials were presented in Dutch. All students received partial credit towards a course requirement. Their average age was 18.99 and 51 of the participants were women, 39 were men, and 1 participant did report gender.

Design

A 3 (type of comparison target: attainable, unattainable, no comparison) $\times$ 2(type of domain: domain match versus domain mismatch) between-participants design was employed. Participants viewed the comparison targets, then they completed all dependent measures.

Procedure

The same cover story and procedures were used as in Study 1. The targets were described as attainable (older) or as unattainable (younger). Their success was attributed to verbal ability. A no-comparison

control condition was also included in the study, in which participants read about an unrelated university project. After reading about the target, participants completed a four-item measure of their current thoughts. Two items were fillers ('I am feeling more tired than usual' and 'Right now, I am worried about the other things that I need to do today'). The other two items measured feelings of threat (alpha = 0.79): 'Right now I feel I have less scholastic ability than others' and 'At this moment, I am concerned that compared to others, I am seen as a failure.' Both threat items were taken from previously published research (Marx, Stapel, & Muller, 2005). Threat measures were followed by measures of self-evaluations. Three-items (alpha = 0.85) 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.'

After completing those measures, participants were presented with the instructions for the RAT. The RAT was framed as a measure of verbal ability or as a measure of logical ability. Following the instructions, participants were asked about their perceptions of the test. First, they were asked to rate on 5-point scales how well they thought they would perform on the test (1 = not at all well, 5 = very well) and how difficult they thought the test would be (1 = not at all difficult, 5 = very difficult). Then participants completed 10 RAT items.

After the test, participants were asked to rate, on a 5-point scale (1 = not at all well, 5 = very well), how well they thought that Hans would do on this type of test. There were also asked how likely it was that they would be as successful as Hans, again on a 5-point scale (1 = not at all likely, 5 = very likely). Participants reported their age and sex, then they were debriefed.

RESULTS AND DISCUSSION

When participants were exposed to unattainable targets, they were expected to feel threatened and report lower self-evaluations. However, when presented with performance opportunities that mismatched the domain of the target, participants were expected to have higher performance expectations and view the test as less difficult. Analyses examining the effects of age and target sex were conducted. Age and sex did not statistically significantly interact with type of target or type of measure, therefore all analyses collapsed across sex and age, $ps > 0.25$.

Perceptions of the Target

To confirm that our measure of attainability was effective, we conducted a 2(type of target: attainable or unattainable) × 2(type of test:match (verbal) or mismatch (analytical)) between-participants ANOVA on attainability ratings. As seen in Table 2, only a main effect of target type on participant estimates of the likelihood of attaining similar success as the comparison target emerged as statistically significant, $F(1, 56) = 14.49$, $p < 0.001$. Those reading about the attainable target felt it was more likely that they would be as successful as the target ($M = 3.23$) than those reading about the unattainable target ($M = 2.21$). Thus, as in the pilot study, the manipulation of comparison target age altered participant views of the comparison targets as attainable. Although attainability was measured after participants completed the performance task (to avoid making comparison with the target conscious and directed), the type of test did not affect perceptions of the comparison target.

We also wanted to ensure that our manipulation of domain matching was effective. To do that, we examined the effects of type of target and type of test on participant estimates of how well the target would do on the test. As expected, only a statistically significant main effect of type of test emerged,

Table 2. Mean (and standard deviations) of participant responses by type of target and type of task (Study 2)

	Unattainable target		Attainable target		No comparison	
	Matching	Mismatch	Matching	Mismatch	Matching	Mismatch
Self-Evaluations*	6.21^{a} (1.05)		9.87^{b} (2.17)		8.03^{b} (1.72)	
Threat*	6.79^{a} (2.21)		5.13^{b} (1.45)		5.32^{b} (1.58)	
Target performance expectation	4.21^{a} (0.58)	3.53^{b} (0.91)	4.44^{a} (0.63)	3.47^{b} (0.64)	—	—
Participant performance expectation	2.57^{a} (0.85)	4.13^{b} (1.19)	3.69^{c} (0.60)	3.40^{c} (0.98)	3.47^{c} (1.13)	3.13^{c} (0.89)
Test difficulty	4.14^{a} (1.03)	2.80^{b} (0.86)	3.38^{c} (1.03)	3.13^{c} (0.92)	3.33^{c} (0.90)	3.50^{c} (0.82)
Performance	7.21^{a} (1.58)	9.33^{b} (2.16)	7.50^{a} (1.46)	7.60^{a} (2.50)	7.13^{a} (2.03)	7.06^{a} (1.53)

Note: Self-evaluations and threat were measured prior to instructions regarding test.
Note: Means not sharing a superscript differ at the 0.05 level or less.

$F(1,56) = 20.59$, $p < 0.001$. Participants expected that Hans would do better on the test when it was described as a measure of verbal ability ($M = 4.33$) than when it was described as a measure of logical ability ($M = 3.50$). Thus, as presented in Table 2, our manipulation of matching was effective.

Self-perceptions

Feelings of Threat

We expected that the unattainable comparison target would lead participants to feel less competent and less successful than the attainable comparison target. Because threat was measured prior to test instructions, a one way ANOVA comparing participants in the attainable, unattainable, and no-comparison condition was used to examine how threatened participants felt. Results are presented in Table 2. This analysis revealed a statistically significant effect of type of target, $F(2,88) = 7.86$, $p = 0.001$. A Tukey's HSD test revealed that participants viewing the unattainable comparison target were more threatened ($M = 6.79$) than those viewing either the attainable target ($M = 5.13$, $p = 0.001$) or the no-comparison condition ($M = 5.32$), $p = 0.005$. The attainable and control conditions did not statistically significantly differ from one another, $p > 0.9$. Thus, exposure to unattainable targets lowered participants' feelings of competency.

Self-evaluations

In addition to threatening feelings of competency, the unattainable comparison target was also expected to negatively impact global self-evaluations. Again, because self-evaluations were measured prior to test instructions, a one way ANOVA was used to examine the effects of target type of self-evaluations. Means and standard deviations are presented in Table 2. This analysis revealed a statistically significant effect of type of target, $F(2,88) = 33.90$, $p < 0.001$. A Tukey's HSD test revealed the expected pattern. Participants viewing the attainable comparison target had higher self-evaluations ($M = 9.87$) than those in the control condition ($M = 8.03$), and in both of these conditions self-evaluations were higher than in the unattainable target condition ($M = 6.21$), all ps < 0.001.

Performance Expectations and Test Difficulty

We expected that how individuals felt about their abilities and likelihoods of success would be influenced by the comparison targets, as shown in the earlier analyses. However, we also expected that these feelings would change when participants received a performance opportunity and that how those feelings would change would depend on how the performance opportunity was described. Among participants who were initially threatened by the comparison target, when they were given an opportunity to perform in a different domain from the comparison target, they were expected to view the task as less difficult and anticipate better performance. That is, a mismatching test was expected to resolve or allay feelings of threat. However, when these same participants were given an opportunity to perform in the same domain as the comparison target, they were expected to see the task as more difficult and have lower performance expectations. To test these hypotheses, a series of 2(type of target: attainable or unattainable) $\times$ 2(type of test: match (verbal) or mismatch (analytical) ANOVA were conducted on performance expectations and estimations of test difficulty.

As expected, only the interaction effect of type of target and type of test statistically significantly affected performance expectations, $F(2,85) = 9.51$, $p < 0.001$. As can be seen in Table 2, further analysis revealed that among participants in the attainable target and control conditions, type of test did not influence expectations. However, among those in the unattainable conditions, participants taking the matching test had lower expectations ($M = 2.57$) than those taking the mismatching test ($M = 4.13$), $F(1,89) = 19.57$, $p < 0.001$.

A similar pattern was found for test difficulty. A statistically significant main effect emerged such that participants felt that the matching test would be more difficult ($M = 3.60$) than the matching test ($M = 3.15$), $F(1,85) = 5.91$, $p = 0.02$. This effect was qualified by the expected statistically significant interaction effect, $F(2,85) = 5.28$, $p = 0.007$. As with performance expectancies, only among participants viewing the unattainable comparison target did the type of test affect perceptions of test difficulty, $F(1,89) = 15.50$, $p < 0.001$. In that condition, the matching test was perceived as more difficult ($M = 4.14$) than the mismatching test ($M = 2.80$).

Performance

As in Study 1, we expected that the type of target and how the test was described would determine how participants performed on that test. As expected, only the interaction effect of type of target and type of test statistically significantly affected performance, $F(2,85) = 3.01$, $p = 0.05$. As can be seen in Table 2, further analysis revealed that among participants in the attainable target and control conditions, type of test did not influence performance. However, among those in the unattainable conditions, participants taking the matching test solved fewer problems ($M = 7.21$) than those taking the mismatching test ($M = 9.33$), $F(1,89) = 9.12$, $p = 0.003$. Thus, it appears that unattainable targets lead to boosts in performance in alternative domains, relative to a control condition.

The Relationship Between Target, Test, Expectancies, and Performance

The model presented here suggests that when participants see an unattainable comparison target, they experience feelings of threat. It also suggests that when participants are presented with a performance opportunity that is from a different domain than that of the comparison target, they experience a rise in performance expectations and, ultimately, perform better. Conversely, the model suggests that when participants are presented with a performance opportunity that is from the same domain as the

Table 3. Regression results showing that the interaction effect of type of target and type of test on performance was mediated by performance expectancies

	Equation 1 (criterion SCORE)		Equation 2 (criterion expected performance)		Equation 3 (criterion SCORE)		
Predictors	*b*	*t*	*b*	*T*	*b*	*t*	ΔR^2
Type of target	−0.267	−2.18*	−0.30	−2.66**	−0.17	−1.34	0.07*
Type of test	0.174	1.42	−0.09	−0.80	0.16	1.17	0.03
Target × Test	−0.243	−1.98*	−0.44	−3.86**	−0.11	−0.77	0.06*
Expected performance					0.335	2.26*	0.10**
Expected performance × Type of test					−0.010	−0.69	0.007
R^2 total	0.16		0.28		0.27		

Note: $^{*}p < 0.05$; $^{**}p < 0.01$.

comparison target, they may become more pessimistic about their possible performance and perform worse. Thus, the model predicts mediated moderation—how the type of comparison target affects performance is moderated by the type of test. The change in participant performance expectations is responsible for that moderation (Muller et al., 2005). To test this model, three regression equations, as outlined by Muller et al., were computed.

In all equations, only participants who viewed a comparison target were included in the analyses. Additionally, all continuous variables were standardized. All relevant betas for all three regression analyses are presented in Table 3.

In Equation 1, the type of comparison target, the type of test and the interaction effect of target and test were regressed on the dependent variable, performance. As shown in the first column of Table 3 (and replicating the earlier analysis of variance), both type of test and the interaction effect predicted performance.

In Equation 2, type of target, type of test, and the interaction effect were regressed on the mediating variable, performance expectations. As shown in the second column of Table 3, type of test and the interaction effect predicted performance expectations. Again, replicating the earlier analysis of variance, the mismatching test was associated with higher performance expectations and this difference was increased when the comparison target was unattainable.

Finally, a third equation predicting performance was computed. In the first step of Equation 3, type of target, type of test, and the interaction effect were entered. In the second step, performance expectations and the interaction of performance expectations and type of test were entered in the equation. As can be seen from the last column in Table 3, Equation 3 reveals a statistically significant effect of expected performance on actual performance. The direct effect of type of target, type of test, and the interaction effect of target and test was less, once the mediator was controlled. A Sobel test (Preacher & Leonardelli, 2001) revealed that the decrease in the direct effect of the interaction effect of the type of target and type of test was statistically significant, $z = 1.94$, $p = 0.05$.

The results of the regression analyses described in Table 3 demonstrate a pattern congruent with mediated moderation (Muller et al., 2005). The effect of type of target on performance expectations is statistically significantly moderated by the type of test. And, when participant performance expectations are included in the model, the interaction effect of the type of test that they are taking and the type of comparison target about which they have read no longer predicts their performance, whereas

performance expectations do. Thus, these data provide strong support for our model in which the changing domains can lead threatening social comparisons to positively impact performance.

STUDY 3

Studies 1 and 2 demonstrated that mismatching of target and performance domain leads to performance benefits for those exposed to unattainable targets while matching does not. However, it also possible that target domain and testing domain need not be orthogonal for performance benefits to occur. Instead, the domains may be described as overlapping, with the target domain encompassing the test domain. In such a case, the domains may be neither a match nor a mismatch. If the target's domain is not specifically the same as the test domain, individuals may see an opportunity to regain self-regard by performing well on the test. Therefore, Study 3 examines the effects of the globality or specificity of target success on performance. In addition, self-evaluations and performance are measured separately, with some participants completing the self-evaluation measure prior to performance, and other participants only completing the measure of performance. Although other studies have included both measures of self-evaluation and performance as within-participant measures, we wanted to exclude the possibility that asking participants to self-evaluate would lead to directed social comparison (Stapel & Suls, 2004), while also examining the effects of global versus specific comparison targets on self-evaluations. One could hypothesize that specific comparison targets should be regarded as less threatening because they only threaten one self-aspect (e.g., one's chemist-self) while globally successful target may threaten many self-aspects. Alternatively, one could expect that globally successful targets may be less threatening because they are ambiguous and one could choose to believe that globally successful targets would not be as good in important and self-relevant domains. Here, we test these hypotheses and suggest that globally successful targets will be evaluated in defensively. Global and attainable targets should result in the highest self-evaluations, while specific and unattainable should result in the lowest self-evaluations.

When the target's success is described as global in nature, unattainable targets are expected to lead to better performance relative to attainable targets. When the target's success is described as specific in nature, this performance benefit is not expected.

Method

Participants

Participants were 146 Dutch university students and all materials were presented in Dutch. All students received partial credit towards a course requirement. Although gender information was not collected, participants were representative of the overall participant pool, which consisted of 70% females. All participants were undergraduate students in their first two years of school.

Design

A 2(type of target: attainable vs. unattainable) $\times$ 2(type of domain: specific vs. general) between-participants design was employed. In the specific target domain conditions, target success was attributed to their verbal ability. In the global target domain conditions, target success was attributed to

general intelligence. In all conditions, the test was described as a measure of verbal ability. Therefore, in the global target conditions, the test and target domains overlapped. In the specific target domains, the test and target domains were identical. Participants viewed the comparison targets, then either completed the performance task, or completed a measure of self-evaluations.

Materials

Self-evaluation Measure

The 3-item self-evaluation measure (alpha = 0.67) 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'.

Procedure

The same cover story and procedures were used as in Studies 1 and 2. After reading about the target, participants were asked to complete a three-item self-evaluation measure or complete the RAT. In all conditions, the RAT was framed as a measure of verbal ability. Therefore, in the global condition the testing domain was encompassed by the domain of target success and overlapped with it, but was not restricted to it. In the specific condition, the testing domain and domain of target success were exactly the same and completely overlapped.

RESULTS AND DISCUSSION

Self-evaluations

A 2(type of target) × 2(target domain) between-participants ANOVA examining self-evaluations was conducted. Specifically described ($M = 6.97$) and globally successful targets ($M = 7.46$) did not lead to reliable differences in self-evaluations. However, participants exposed to the attainable comparison targets ($M = 11.3$) reported higher self-evaluations than those exposed to the unattainable comparison targets ($M = 10.13$), $F(1,66) = 4.86$, $p = 0.03$.

Although the interaction effect of target status and target domain was not statistically significant, specific hypotheses were made and exploratory tests of these hypotheses were conducted. Contrast analyses supported the hypothesis that global and attainable comparison targets would lead to the highest self-evaluations ($F(1,66) = 4.26$, $p = 0.05$), but failed to support the hypothesis that specific and unattainable targets would lead to the lowest self-evaluations, $F < 1.0$. When the mean for participants who read the global and attainable comparison target was contrasted with the mean participant ratings for the other three experimental conditions, their mean self-evaluations were reliably higher.

Performance

A 2(type of target) × 2(target domain) between-participants ANOVA examining performance revealed no statistically significant main effects. However, a statistically significant interaction effect of target

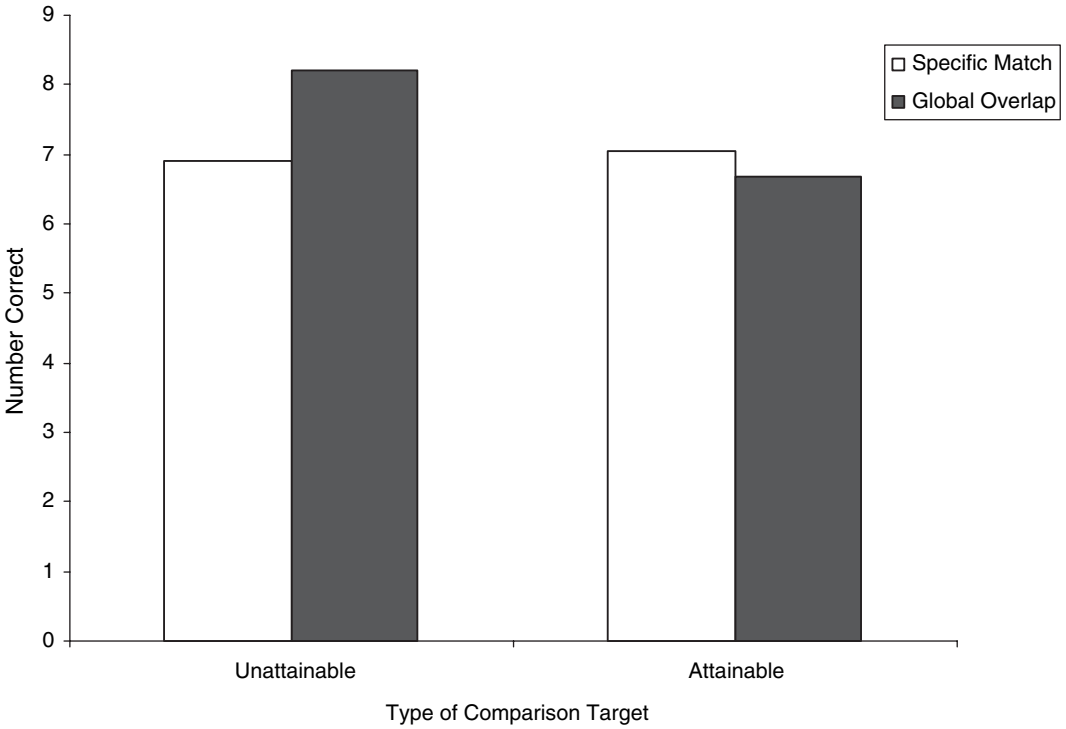

Figure 2. Number correct by role model attainability and overlap of domain (Study 3)

status and target domain on performance was found, $F(1,72) = 5.15$, $p = 0.03$. Among participants who were exposed to the attainable comparison targets, the domain of the comparison target did not affect performance ($M_{\text{specific}} = 7.06$; $M_{\text{global}} = 6.67$) $F < 1$. Among participants who were exposed to the unattainable comparison targets, those exposed to the specific comparison target performed worse ($M = 6.90$) than those who exposed to the global comparison target ($M = 8.21$), $F(1,74) = 5.97$, $p = 0.02$ (See Figure 2).

The results of Study 3 supported our hypothesis that when the successes of the targets were attributed to general traits or global abilities, participants would respond as though the target domain of success and the performance domain were mismatching. When the successes of targets were attributed to a general set of traits, unattainable targets led to relatively increased performances. Yet, when the successes of the targets were attributed to a specific trait and matched the domain of the test, no statistically significant differences in performance were found. This suggested that globally successful others may have the same effects on performance as successful others in completed separate domains.

Study 3 replicated Study 2 in showing that the domain of success of the comparison target determines how social comparison impacts self-evaluations: unattainable comparison targets led to lower self-evaluations compared to attainable comparison targets regardless of how the target's success was described. It also showed that for those reading about attainable comparison targets, the domain of the target's success is not relevant to performance. Participants performed equally well, regardless of how that target's success was described and regardless of the relationship between the target's domain and the test domain.

GENERAL DISCUSSION

There are a number of methods for coping with self-evaluation threats elicited by exposure to superior others. One such method is increased performance (see Tesser, 1988). Here we demonstrated that increased performance in response to self-evaluation threats occurs under particular conditions. Specifically, Studies 1 and 2 demonstrated that when individuals are exposed to a demoralizing comparison target, they respond by increasing performance in domains that mismatch with the target's performance domain. Similarly, participants' perceptions of the task were affected by target attainability and the task domain. When targets were unattainable and the task from the target's domain, participants anticipated that the task would be more difficult, were less satisfied with their performance, and predicted that they would do worse than when the targets were attainable or the task was from another domain.

Study 2 also provided evidence as to the mechanism by which changing domains allows for threatening comparison targets to positively influence performance. In that study, when participants exposed to the unattainable comparison target received an opportunity to perform in a different domain than the target, their performance expectancies rose and their performances improved. Thus, changing domains transformed a fear of inadequacy into an expectation of success.

Study 3 found that the target and testing domains need not be mutually exclusive in increase performance. When the target domain was global and the testing domain specific, the unattainable target led to better performance relative to the attainable target.

In the present research, we demonstrated that an unattainable comparison target leads to increases in performance relative to that of an attainable target. We concluded that self-evaluation threats posed by unattainable role models lead to the increased performance in alternative domains.

Implications

We suggest that more successful others should have a positive influence on individuals' performances in domains other than that of the comparison target. Several other models predict that upward social comparison should lead to better performance. For example, the reflection-evaluation model (Markman & McMullen, 2003) argues that evaluation of the self against an upward comparison target should initiate a series of processes culminating in better performance on some tasks and Seta (1982; Seta, Seta, & Donaldson, 1991) has shown that a successful co-actor can lead to increases in performance. However, these studies do not fully take into account the relationship between the perceiver's task and the target's success or the relationship between the perceiver and the comparison target. Both are necessary in understanding how comparison with successful others influences performance, as the present studies demonstrated. Thus, the current studies do not contradict previous work, but extend that work.

Finally, Blanton et al. (1999) have suggested that an individual's choice to compare upward may benefit their performance. In their research, students who chose to compare themselves with more successful students did better academically. Like other social comparison research, Blanton et al.'s work examined chosen comparisons. In contrast, the current studies are more concerned with the comparisons that may occur in the absence of intention or awareness. A number of studies have suggested that social comparison may be an automatic process, occurring spontaneously (Mussweiler, Ruter, & Epstude, 2004; Stapel & Blanton, 2004; Stapel & Suls, 2004) and regardless of the characteristics of the target (Gilbert, Giesler, & Morris, 1995). Again, the current studies do not

contradict previous research, but illuminate the *downstream* impact upward social comparison, focusing on both self-evaluations *and* performance.

Limitations and Future Research

In the current studies, an ambiguous situation was created by describing the target's domain of success as global and the performance domain as specific. We suggest that participants respond to the ambiguous situation as if the domains were mismatching, and interpreted the performance task as a means of repairing and maintaining positive self-evaluations. This has a number of implications. First, it contributes to the growing literature showing that people interpret information in social comparison situations in defensive ways (Schwinghammer, Stapel, & Blanton, 2006). Second, it suggests that unattainable upward comparisons can influence performance in many different situations; whenever and wherever individuals are exposed to successful others. However, before such conclusions may be drawn, additional research is necessary. To fully understand the generalizability of the relationship between unattainable comparison and performance, future research must also examine what individuals perceive to be the relationship between globally and specifically defined domains and what the boundary conditions are to the perception that such domains mismatch.

The current studies sought to create situations in which spontaneous comparison occurred— in which individuals were not directly asked to compare themselves to another but spontaneously engaged in comparison following exposure. Recent research has suggested that whether or not individuals are explicitly asked to compare themselves to others can lead to distinctly different consequences (Stapel & Suls, 2004). In many situations, such as when employers are attempting on motivate employees or parents their children, individuals are asked to compare themselves to more successful others as a means of creating motivation. We would argue that in such situations, being directed to compare oneself to another should lead to similar feelings of threat found in these studies, perhaps even magnified levels of threat. Simply being told that others are more successful and that one should follow their examples may be considered to be threatening. However, boundary conditions such as beliefs about the malleability of domain, state of mind when encountering such upward comparison targets, and self-construal (Stapel & Koomen, 2001) may determine when and where upward social comparisons impact performance.

Conclusion

As we stride across university campuses, we see more junior and more senior colleagues who have achieved success in our field, in fields distinct from ours, and in arenas that encompass our own ambitions. This research suggests that those who make us feel best about ourselves, and lift our heads with pride, may not be the ones that lead us to perform our best. Ironically, it appears that those colleagues that most demoralize us, and are not in our specific field may provide the greatest long-term benefit.

ACKNOWLEDGEMENTS

This publication was supported in part by grant number T32 MH19728 from the National Institute of Mental Health, and National Research Service Award F31 MH64238-01 awarded to

Camille S. Johnson. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the National Institute of Mental Health. This publication was also supported in part by a grant 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.

REFERENCES

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.

Crocker, J., & Major, B. (1989). Social stigma and self-esteem: The self-protective properties of stigma. *Psychological Review, 96*, 608–630.

Festinger, L. (1954). A theory of social comparison processes. *Human Relations, 7*, 117–140.

Gilbert, D. T., Giesler, R. B., & Morris, K. A. (1995). When comparisons arise. *Journal of Personality and Social Psychology, 69*, 227–236.

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.

Markman, K. D., & McMullen, M. N. (2003). A reflection and evaluation model of comparative thinking. *Personality & Social Psychology Review, 7*, 244–267.

Marx, D. M., & Roman, J. S. (2002). Female role models: Protecting women's math test performance. *Personality and Social Psychology Bulletin, 28*, 1183–1193.

Marx, D. M., Stapel, D. A., & Muller, D. (2005). We can do it: The interplay of construal orientation and social comparisons under threat. *Journal of Personality and Social Psychology, 88*, 432–446.

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.

Muller, D., Judd, C. M., & Yzerbyt, V. Y. (2005). When moderation is mediated and mediation is moderated. *Journal of Personality and Social Psychology, 89*(6), 852–863.

Mussweiler, T., Ruter, K., & Epstude, K. (2004). The man who wasn't there: Subliminal social comparison standards influence self-evaluation. *Journal of Experimental Social Psychology, 40*, 689–696.

Preacher, K. J., & Leonardelli, G. J. (2001, March). Calculation for the Sobel test: An interactive calculation tool for mediation tests [Computer software]. Available from http://www.quantpsy.org

Schwinghammer, S. A., Stapel, D. A., & Blanton, H. (2006). Different selves have different effects: Self-activation and defensive social comparison. *Personality and Social Psychology Bulletin, 32*, 27–39.

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.

Stapel, D. A., & Blanton, H. (2004). From seeing to being: Subliminal social comparisons affect implicit and explicit self-evaluations. *Journal of Personality and Social Psychology, 87*, 468–481.

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., & 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. (1997). A threat in the air: How stereotypes shape intellectual identity and performance. *American Psychologist, 52*, 613–629.

Tesser, A. (1988). Toward a self-evaluation maintenance model of social behavior. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 20, pp. 181–227). New York: Academic Press.

Wood, J. V. (1989). Theory and research concerning social comparison of personal attributes. *Journal of Personality and Social Psychology, 106*(3), 231–248.

Wood, J. V. (1996). What is social comparison and how should we study it? *Personality and Social Psychology Bulletin, 22*, 520–537.

