

Social Identity versus Reference Frame Comparisons: The Moderating Role of Stereotype Endorsement

Hart Blanton, Charlene Christie, and Maureen Dye

State University of New York at Albany

Received March 29, 2001; revised September 17, 2001; accepted September 19, 2001; published online February 11, 2002

Two studies tested the prediction that belief in a negative stereotype about an in-group will cause members to shift from viewing their in-group as a social identity to viewing it as a frame of reference. The stereotype that was the focus of inquiry was the belief that women have less aptitude at math and spatial tasks than do men. In both studies, female participants took a test of math and spatial ability and then received social comparison information about their abilities relative to a male and a female confederate. In Study 1, participants felt enhanced when the two women outperformed the male confederate, even when this meant that the participants themselves performed worse than the other woman. If participants were first reminded of the negative stereotype, however, they felt best when they outperformed the other woman, even if this meant that the two women performed worse than the man. Study 2 showed that the effects of stereotype activation were especially pronounced among female participants who showed moderate to high levels of stereotype endorsement. These findings suggest that belief in stereotypes about the in-group can lead to in-group comparison and contrast, even in contexts in which a group member's ability level challenges the validity of the stereotype. © 2002 Elsevier Science (USA)

Key Words: assimilation; contrast; gender; self-esteem; social comparison; social identity; reference groups; stereotypes.

The social psychological literature points to two ways in which gender might influence the social comparison of abilities. We label these two theoretical perspectives the *social identity view* and the *reference frame view*. First, social identity theory (Tajfel, 1978; Tajfel & Turner, 1979) and its extension, self-categorization theory (Turner, Hogg, Oakes, Reicher, & Wetherell, 1987; Turner, Oakes, Haslam, Haslam, & McGarty, 1994), indicate that increased gender salience might heighten concern for the maintenance of a positive gender identity. When this occurs, individuals will want their “in-group gender” to be positively distinct from their “out-group gender,” and they will view their own self-attributes as psychologically interchangeable with the attributes of other in-group members (Turner, 1985). As a result, upward same-gender social comparisons and downward cross-gender social comparisons should have the potential to raise self-esteem, whereas downward same-gender and upward cross-gender social comparisons should have

the potential to lower self-esteem (Blanton, Crocker, & Miller, 2000; see also Blanton, 2001). In contrast to this theoretical view, traditional social comparison theory has considered how people treat their gender as a reference frame, against which their personal abilities are then evaluated. This literature indicates that increased gender salience might heighten concern for doing well in relation to same-gender standards of comparison and simultaneously lower concern for doing well in relation to cross-gender standards of comparison. When this occurs, individuals will view same-gender targets as more relevant standards of comparison than cross-gender targets (e.g., Feldman & Ruble, 1981; Major & Forcey, 1985; Miller, 1984; Suls, Gaes, & Gastorf, 1979). As a result, upward social comparisons should threaten self-esteem, and downward social comparisons should enhance self-esteem to greater degrees, when they are made with same-gender targets of comparison than when they are made with cross-gender targets of comparison (Major, Sciacchitano, & Crocker, 1993; see also Blanton, 2001).

The current studies test the prediction that negative gender stereotypes can alter the manner in which gender influences the social comparison of abilities. Specifically, we predict that the negative stereotype regarding women's math and spatial abilities, when activated, will cause women

The authors thank Adam Galinsky, Jim Jaccard, and Mark Muraven for their comments and suggestions. Thanks especially go to Alexander Haslam for providing a signed review that improved the Discussion section and to Judith M. Harackiewicz for a thoughtful and speedy review.

Address correspondence and reprint requests to Hart Blanton, Department of Psychology, State University of New York at Albany, Albany, NY 12222. E-mail: hblanton@albany.edu.

to shift from treating their gender as a social identity to treating it as a reference frame. The rationale for this prediction is based on the social construction of the stereotype regarding women's math and spatial abilities. We argue that women often show some degree of endorsement of these stereotypes. As a result, they will restrict their definitions of "high" and "low" abilities in these domains to same-gender norms when the stereotypes are activated, which will then cause them to use same-gender standards of comparison when evaluating their own abilities. Moreover, we predict that stereotype activation will cause gender to be treated as a reference frame, even in social contexts where it would be treated as a social identity if stereotypes were not activated.

Endorsement of the Gender and Math Stereotype

Many have noted that the psychological dynamics surrounding gender stereotypes are not typical of the dynamics that surround many other social stereotypes. It appears that women, like men, often have what might be termed a "cooperative relationship" with the stereotypes about their group. As Eberhardt and Fiske (1994) have noted, romantic relationships create incentives for heterosexual men and women to accept stereotypical gender roles (see also Eagly, 1987; Peplau, 1983). Reflective of this, gender stereotypes appear to reinforce social constructions of men and women as two different "groups" that possess complementary strengths and weaknesses (Glick & Fiske, 1996). By comparison, constructions of Blacks and Whites often emphasize antagonistic relationships. The result is that women at times view their diminished status relative to men as legitimate, whereas African Americans more typically view their diminished status relative to Whites as illegitimate (see Major, 1994).

One consequence of legitimating status differences is endorsement of in-group stereotypes, even when they are negative (see Jost & Banaji, 1994). As a specific case, it appears that women show some acceptance of the stereotype that their gender possesses less aptitude for solving math and spatial problems than does the male gender. There is strong developmental evidence, for instance, that parents' gender stereotypical beliefs undermine girls' belief in their math abilities (Eccles, Jacobs, & Harold, 1990; Frome & Eccles, 1998; Jacobs & Eccles, 1992), and there is evidence that socialization contexts that increase the salience of young girls' gender diminishes their aspirations in male-stereotypical ability domains (Abrams, Sparkes, & Hogg, 1985; Abrams, Thomas, & Hogg, 1990). These findings suggest that gender socialization often undermines women's confidence in their math and spatial abilities and that this can flow from informal pressures to adopt gender stereotypes. Admittedly, however, these findings fall short of showing that women often take the next step and internalize beliefs that the stereotypes about their abilities are true. For instance, women may be more likely than men to believe

that they have poor math abilities but still reject the stereotype that women *as a group* have less math ability than do men *as a group*.

There is reason to believe, however, that women will often show signs of stereotype endorsement. Laboratory studies suggest, for instance, that young women have at least implicit belief in gender stereotypes. Various research projects have shown that the increased salience of gender stereotypes undermines women's performance on math tests (e.g., Gonzales, Blanton & Williams, 2001; Inzlicht & Ben-Zeev, 2000; Spencer, Steele, & Quinn, 1999). This suggests that women often possess some degree of internalization of the stereotypical view, even if they do not explicitly endorse the stereotype as true. It seems likely, however, that many will take the next step as well and explicitly endorse gender stereotypes. In fact, some theoretical perspectives within the psychological community have suggested that women *should* believe the stereotype related to their abilities. This is because some psychological researchers have argued that gender differences in ability do exist and that they reflect biological differences in aptitude (e.g., Benbow & Stanley, 1983).

Although most women will not be familiar with the specific reports that take this view, many appear to believe that "research has shown" that women are less able than men in these domains. We became aware of this while running pilot studies that lead to the studies presented here (see Blanton, 2001). In an initial study, female participants interacted with a male experimenter who made an offhand comment that he did not think women were as capable as men at tasks requiring math and spatial ability. When asked during the debriefing whether they had been offended by this statement, many participants stated that they had not been because they had assumed that the statement was accurate. This anecdote certainly is not a systematic demonstration that women often buy into the stereotypes about their math and spatial abilities, but the passive acceptance we observed in this pilot work stands in stark contrast to the response we might expect from members of many other stereotyped groups.

In summary, there is good reason to believe that women often show some degree of belief in or endorsement of the stereotypes that women have worse math and spatial abilities than do men. Although the belief in these in-group stereotypes might not be high in absolute terms, it seems likely that it is higher than would be found for most targets of negative stereotypes. To the extent that women buy into negative gender stereotypes, this should have dramatic consequences for social comparisons of ability.

Stereotype Endorsement and Social Comparison

If a negative ability stereotype about one's own group is true, then in-group members are more relevant standards of comparison than are members of the advantaged out-group

(Goethals & Darley, 1977). Indirect evidence that activation of gender stereotypes will narrow comparisons in this way can be found in person perception research by Biernat and Manis (1994; see also Biernat, Manis, & Nelson, 1991). According to their "shifting standards model," group members are typically evaluated by others on stereotype-relevant dimensions against in-group standards. Thus, perceivers who hold the stereotype "women are less competent than men" will typically evaluate women based on their ability to do well relative to other women and men based on their ability to do well relative to other men (Biernat & Manis, 1994, Study 1). In a similar fashion, we are suggesting that women who believe the stereotype "women are less competent than men at math" will evaluate their *own* math ability relative to the math ability of other women—not the abilities of other men. Moreover, we are predicting that activation of the math stereotype will cause this comparison style, even in contexts that would otherwise promote the treatment of gender as a social identity.

The competing hypotheses are that activation of a negative stereotype will either increase the tendency to view gender as a social identity or leave such tendencies unaffected. In a recent study, Blanton et al. (2000) found that Black students showed higher self-esteem following upward as opposed to downward social comparison with another Black student. Although the negative stereotype about Blacks' intellectual abilities was not explicitly activated, the testing situation contained aspects of stereotype threat that should have heightened the salience of negative stereotypes (Steele & Aronson, 1995). It is possible that these stereotypes were responsible for Black students identifying with one another's performance. Such effects of stereotypes are not predicted in Study 1, however, because female participants are expected to show a greater acceptance of the negative intellectual stereotypes about their gender than Black participants typically show for the negative intellectual stereotypes about their race.

Creating an Intragroup/Intergroup Dilemma

The current studies present female participants with comparison information that should increase the salience of intergender (social identity) comparisons. In evidence of this, we predict that participants will feel enhanced by the superior performance of women relative to men and diminished by the superior performance of men relative to women. This comparison situation will be set up in such a way, however, that *the status of women relative to men improves at the cost of the participant's own status relative to other women*. As a result, any manipulation that shifts attention away from the intergender implications of the comparison information and toward the intragender implications of the comparison information should result in a shift in the contingencies surrounding self-evaluation. When a manipulation causes such a shift, participants

should feel diminished by the comparison information that was previously enhancing and enhanced by the comparison information that was previously diminishing. This logic provides the tool we need to determine whether stereotypes about women's math and spatial abilities will cause female participants to shift from viewing their gender as a social identity to viewing it as a reference frame.

STUDY 1

Method

Participants. A total of 61 female college students from introductory psychology classes participated for course credit. Of these participants, 61% were White, 23% were Asian American, 8% were African American, 2% were Latino, and the remaining indicated a multiracial or multiethnic identification. The mean age of the sample was 19 years.

Social comparison procedure. The social comparison manipulation was adapted from a study by Sanders, Gastorf, and Mullen (1979). Participants in all conditions took a test of math and spatial abilities and received identical feedback about their performance. Following this, they were presented with one of two types of social comparison information about both a male and a female confederate. In the *women-up* condition, each participant learned that she and another woman had outperformed a male but that she personally had performed worse than the other woman. An alternative way of stating this is to say that participants were exposed to downward social comparison information regarding a male confederate's performance and to upward social comparison information regarding a female confederate's performance. In the *women-down* condition, each participant learned that she and another woman had performed worse than a male but that she personally had outperformed the other woman. Or, participants were exposed to upward social comparison regarding a male confederate's performance and to downward social comparison regarding a female confederate's performance.

In the absence of a negative stereotype, we predicted higher self-evaluations following social comparison in the *women-up* condition than in the *women-down* condition. The rationale for this was based on Turner's (1985) meta-contrast principle. We predicted that our comparison procedure, which simultaneously presents both same-sex and cross-sex comparison information, would heighten the intergroup implications of the comparison information and thereby heighten the desire for positive distinctions between the in-group gender and the out-group gender. As a result, we expected that women would feel best when they outperformed the male confederate.¹ Support for this prediction

¹ The conditions created by this procedure have conceptual similarity to

can be found in a study by Sanders et al. (1979). They gave both male and female participants comparison information about both same-sex and cross-sex targets of comparison and then manipulated the relative performance of male versus female targets of comparison. Results showed that participants cued in to the relative standing of same-sex versus cross-sex targets and that they made their highest ability ratings when same-sex targets outperformed cross-sex targets. This finding suggests that study participants will spontaneously view gender as a social identity if they are presented with feedback about both male and female targets of comparison. The current study attempted to replicate the sort of social identity effect shown in Sanders et al. to determine whether the activation of a negative stereotype would shift participants from viewing their gender as a social identity to viewing it as a frame of reference.

Design. Participants were randomly assigned (with cell sizes ranging from 14 to 16) to one of the four experimental conditions based on a 2 (Stereotype Salience: low vs high) \times 2 (Social Comparison: women-up vs women-down) between-subjects factorial design. The primary dependent variable was self-ratings of ability after task performance.

Procedure. The procedure in the current study was adapted from Blanton et al. (2000). Each participant was met by a female experimenter, who asked her to wait for two other participants. The other participants, actually a male and a female confederate, arrived after the participants had an opportunity to sign the experimental consent form. At this point, the experimenter directed all three individuals to a laboratory, where she explained that the purpose of the study was to “standardize” a new measure of “natural math and spatial ability.” She explained that this measure was unlike tests they may have taken in the past. Although the test was based on mathematical and spatial reasoning ability, she said, it was not thought to be contingent on previous level of schooling in math and spatial reasoning. The advantage of this instruction set is that it should minimize stereotypical gender expectations for the task (see Spencer et al., 1999). Participants were then told that most college students could solve these problems regardless of their level of training and that their own scores would be based on a combination of their speed and accuracy. The

those that have been of interest in research on the “frog pond effect” (Davis, 1966). The frog pond effect proposes that individuals evaluate personal ability relative to in-groups and pay little attention to out-groups. The prediction we make are the reverse of this. We predict that self-evaluations will be based on intergroup standing, with little concern for in-group standing. The rationale for our prediction is that participants observe their in-group’s performance in the context of salient feedback about the out-group’s performance. This is different from what occurs with the frog pond effect. In this research, individuals are aware of in-group performance levels, but they are mostly unaware of performance levels in more distant out-groups (see Marsh & Parker, 1984; McFarland & Buehler, 1995).

advantage of linking performance to reaction times for an unfamiliar task is that this made it difficult for participants to determine how they had done, increasing the likelihood that they would accept the test feedback when they received it.

For those in the high stereotype salience condition, the experimenter introduced the notion that women might perform worse than men, even on this new task. She stated,

The point of this study is that we are hoping to standardize this test. What do we mean by standardizing? By standardizing, we hope to determine if different groups perform differently. For instance, some of the men working on this project believe that women will do poorly relative to men, and your performance on this test will help us determine if this is true.

For those in the low salience condition, the experimenter made no mention of gender, although she did mention that the test was being standardized. She stated,

The point of this study is that we are hoping to standardize this test. What do we mean by standardizing? By standardizing, we hope to determine how most people perform on this test. For instance, since you three are college students, you can thus help us know how college students typically perform.

Participants then took the test on the computer in an isolated terminal. This program presented a straightforward set of computational and spatial problems, with answers presented in a multiple-choice response format. As examples, the test asked participants, “What is the remainder after 72 is divided by 3?” and “When a cube is sitting on a table, what is the maximum number of sides one can view at once?” When participants finished, the computer delivered feedback in the form of speed and accuracy scores, but it did not provide them with any information about what these scores might mean. Instead, it instructed them to meet the experimenter in a neighboring room, where she would help them to code their scores. In all conditions, the participant was the first to return to the room, where she was told that she had scored a 67 out of 100. The upward and downward social comparison targets then entered the room. The participant thus overheard an upward comparison target of comparison receiving a 99 out of 100 and then a downward target of comparison receiving a 35 out of 100. In half of the conditions (women-up), the upward comparison target was the female confederate and the downward comparison target was the male confederate. In the other half of the conditions (women-down), the upward comparison target was the male confederate and the downward comparison target was the female confederate. After receiving the performance feedback and the social comparison information, participants then completed an ability appraisal. Following this, they were given a full process debriefing.

Ability evaluation. The primary dependent variable was a seven-item ability appraisal adapted from the performance subscale of the Heatherton and Polivy (1991). Participants rated their agreement with statements describing their per-

Social Comparison X Stereotype Salience

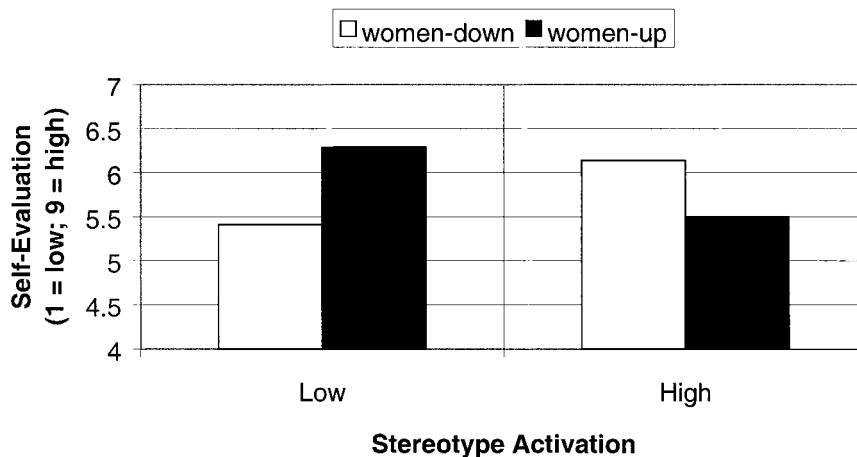


FIG. 1. Ability evaluations in Study 1 as a function of social comparison and stereotype salience.

formance on the task and their competence in general (e.g., “I did well on this task,” “I feel confident about my abilities,” “I feel as smart as others”). Agreement was made using a 10-point Likert scale, with end points of 1 = *strongly disagree* and 9 = *strongly agree* (Cronbach’s alpha = .80).

Predictions. We predicted a crossover interaction between stereotype salience and social comparison. When the stereotype was not activated, we predicted higher ability ratings in the women-up condition than in the women-down condition. This would suggest that the comparison information led women to view their own gender as a social identity, such that their self-evaluations were highest when the two women outperformed the male. When the stereotype was activated, however, we predicted higher ability ratings in the women-down condition than in the women-up condition. This would suggest that the stereotype caused women to view their gender as a frame of reference, such that their self-evaluations were highest when they performed well relative to the other woman.

Results

To test for the predicted interaction, we conducted a 2 (Stereotype Salience) \times 2 (Social Comparison) analysis of variance (ANOVA) on ability evaluation. As predicted, and as shown in Fig. 1, a significant first-order interaction was revealed, $F(1, 57) = 6.37, p < .02$. When the stereotype was not activated, participants reported higher ability in the women-up condition ($M = 6.29, SD = 1.02$), in which the two women outperformed the one man, than in the women-down condition ($M = 5.41, SD = 1.06$), $F(1, 57) = 4.23, p < .04$. When the stereotype was activated, participants reported higher ability in the women-down condition ($M = 6.14, SD = 1.32$) than in the women-up condition ($M = 5.50, SD = 1.28$). Thus, they felt better when they per-

formed better than the other woman, even though this meant that the two women performed worse than the one man. This simple main effect did not reach statistical significance, $F(1, 57) = 2.27, p < .14$, but the significant first-order interaction showed that it represented a significant attenuation of the social identity effect occurring when the stereotype was not activated.

Discussion

As predicted, participants appeared most concerned with the intergroup implications of the comparison information when no mention was made of math and spatial ability stereotypes. The evidence for this is that female participants had higher ability evaluations when the social comparison information indicated that the two women outperformed a male on a test of math and spatial ability. It is important to note that this effect occurred even though participants’ *absolute* performances were held at a constant across conditions. Specifically, participants in all conditions received a 67 out of 100 on the test. In addition, this effect occurred even though participants’ *relative* performances were held at a constant across conditions. Specifically, participants in all conditions performed 27 points higher than a downward comparison other (who scored a 35 out of 100) and 27 points lower than an upward comparison (who scored a 99 out of 100). It is thus safe to conclude that the differences in self-evaluation occurred as a result of the rank ordering of the male versus female confederate. When the stereotype was not mentioned, self-evaluations were higher when the two women outperformed the man than when the man outperformed the two women. This pattern suggests a concern for the maintenance of a positive social identity.

The effect of the comparison information changed dramatically when the negative gender stereotype was activated. The significant first-order interaction indicated that

the previous effect was attenuated by the stereotype activation procedure. Moreover, there was a tendency for participants to have higher ability ratings when the two women performed worse than the man. One might predict that the presence of the negative stereotype would cause women to feel *more* enhanced by the superior performance of other women (e.g., Blanton et al., 2000). Consistent with the notion that the gender and math stereotype is endorsed to some degree by women, however, participants instead appeared to be more focused on their own standing relative to other women. Specifically, they showed a tendency to have more positive self-evaluations when they outperformed the other woman than when they did not.

Although the results lent support to the predictions, there were two primary shortcomings that Study 2 sought to address. First, the results in the stereotype activation condition did not reach statistical significance. Although stereotype activation significantly attenuated the predicted social identity effect and shifted findings in the direction of the predicted reference frame effect, it did not produce a reference frame effect that reached statistical significance using the targeted sample size. Second, the study did not establish that the predicted shift occurred as a result of stereotype endorsement. Study 2 sought to address each of these concerns by determining whether stereotype endorsement moderated reactions to the social comparison feedback. In this way, it was able to determine whether stereotype endorsement was responsible for the observed shift and whether this shift would reach statistical significance among a subsample of women who showed especially high levels of stereotype endorsement.

STUDY 2

Overview

Female participants first completed a questionnaire assessing the extent to which they believed the gender and math stereotype. They then took the measure of natural math and spatial ability under conditions of high stereotype salience and were given social comparison feedback, following the procedure in Study 1. The primary prediction was that level of endorsement would moderate the effect of comparison information on self-evaluations. To the extent participants believed that women have worse math and spatial ability than men, they should show higher ability evaluations when they outperform the other member of their in-group (women-down) than when they do not (women-up). To the extent that the stereotype is not believed, participants should show higher ability evaluations when their in-group outperforms the out-group (women-up) than when it does not (women-down).

The predicted moderating effect of stereotype endorsement was tested against two plausible alternative moderators. It seemed reasonable that the stereotype altered re-

sponses not because it was believed but rather because it changed the emotional reaction to the comparison information. For instance, women may have felt offended by mention of the negative stereotype. As a result, those in the women-up condition may have felt that they had let the other woman down when they had done worse than her on the task. Alternatively, participants may have wanted to see the stereotype discredited when it was mentioned but taken it on themselves to be the one doing the discrediting. As a result, those in the women-up condition may have felt frustrated when the other woman—and not them—was the one to discredit the stereotype. Either of these effects could cause stereotype activation to heighten concern for doing well relative to the other woman, but each would reflect a heightened concern among participants for the maintenance of a positive social identity. We thus developed two measures to test for these alternative moderators. The first measure assessed the degree to which participants were offended by the negative stereotype (stereotype offense), and the second one assessed the degree to which participants took it on themselves to challenge the negative stereotype (stereotype personalizing). Both measures, along with stereotype endorsement measure, are described below. All three ratings were tested as potential moderators of the reaction to social comparison information in the context of the gender and math stereotype.

Participants

A total of 196 female students participated for credit in introductory psychology. Of these, 68% were White, 8% were African American, 5% were Asian American, 5% were Latino, 10% indicated a multiple-group identification, and the remaining students did not respond to the question. The mean age of the sample was 19 years. Ages ranged from 17 to 50, with less than 3% of the sample having ages over 23. Results did not differ if these older individuals were eliminated from the analyses.

Design

Participants were randomly assigned (with cell sizes of $n = 98$) to one of two (Social Comparison Information: women-up vs women-down) experimental conditions. Stereotype endorsement was treated as the focal moderator to determine whether it altered the manner in which comparison direction influenced ability evaluation in the context of the gender and math stereotype.

Procedure

Participants were run in groups of 3 to 10 using a bogus two-study procedure. They first arrived at a laboratory room, where they completed a set of questionnaires that assessed stereotype endorsement, stereotype offense, and

stereotype personalizing. To reduce experimenter demand, these three questionnaires were embedded within a larger set of questionnaires that took about 20 min to complete. Once participants were done, they were directed to a second laboratory on the floor above the first study. There, they were met by a woman who told them that they were going to take a math test on the computer. Participants were randomly assigned either to sit in front of a computer that gave them women-up social comparison feedback or to sit in front of a computer that gave them women-down social comparison feedback.

In all conditions, the computer program first had participants list their age, their year in school, and their gender. It then explained that they were about to take a measure of "natural math and spatial ability," using the same description as in Study 1. To increase stereotype salience, it further told them, "A number of researchers at this university have predicted that women will not perform as well as men on this measure. . . . We are now in the process of standardizing to investigate for the presence of this gender difference." Participants then took the same math test that had been administered in the first study. Afterward, the computer told them that two other students who were their same age and year in school had taken the test on that computer. It then gave them the scores of these students, along with feedback about their own performance. Participants in all conditions learned that they had scored an 82 out of 100. This performance feedback was higher than that which was given in the first study. Participants in Study 1 were told that they scored a 67. During debriefings, however, they often reported that this sounded like a poor score to them. By raising the score to an 82, we hoped to eliminate esteem threat as a potential artifact in our result. Anecdotally, a number of Study 2 participants reported during debriefings that they had been pleased with their own performances. Following this, they then heard that the scores for two other students were a 98 and a 66. Thus, one student scored 16 points worse than they had performed and one student scored 16 points higher than they had performed. When the computer presented the scores of the other two students, it listed their age, their year in school, their gender, and their test scores. This made it possible to manipulate the gender of the upward versus downward social comparison target, as was done in Study 1. It also ensured that the targets of comparison had some degree of comparability with the participant. Once participants viewed the social comparison information, the computer directed them to complete a questionnaire sitting in a folder next to the computer. The seven-item ability evaluation index from Study 1 was included in this questionnaire (Cronbach's $\alpha = .79$).

Measures

Stereotype endorsement. In the initial packet of questionnaires, participants were asked to rate the degree to

which different stereotypes were true. This idea was introduced with the statement,

Below, we have listed some common gender stereotypes. When you read these, you may feel that some are based on gender differences that really do exist. Or, you may feel that the stereotype has no basis in fact. Please read each of the following stereotypes and rate *the degree to which you feel the stereotype is based on true gender differences*. Please answer each question openly and honestly.

They then read a list of 12 stereotypes. Embedded in this list of 12 stereotypes were 2 statements related to the math and spatial ability stereotype. Participants rated their belief in these stereotypes by answering the questions "How much truth is there to the stereotype that 'Men typically have better math skills than women'?" and "How much truth is there to the stereotype that 'Men typically have better spatial skills than women'?" Responses were made on a unipolar 11-point rating scale ranging from 0 to 10. Underneath the numbers 0, 3, 6, and 9 were centered the labels *not at all true*, *some truth*, *a great deal of truth*, and *completely true*, respectively. Responses on these two ratings, $r(194) = .51$, $p < .001$, were then averaged together as a measure of endorsement of the math ability stereotype. For comparison purposes, a measure of general stereotype endorsement was also created by taking an average of all 12 stereotypes. The remaining items tapped beliefs that men are more violent, are more impulsive, are stronger, are more athletic, have better analytic reasoning skills, and have greater intellectual ability, whereas women are more interpersonally sensitive, have better verbal ability, want more children, and are more emotional (Cronbach's $\alpha = .87$).

Stereotype offense. Participants were asked to consider the following:

When another person expresses a stereotypic belief, we at times will get offended. At other times, we may feel less offended. Please rate each stereotype below for how offensive you find it.

Participants then rated how offensive they found each of the 12 stereotypes, using a unipolar 11-point rating scale ranging from 0 to 10. Underneath the numbers 0, 3, 6, and 9 were centered the labels *not at all offensive*, *somewhat offensive*, *very offensive*, and *extremely offensive*, respectively. As before, we computed both the average offensiveness rating for the math and spatial stereotypes, $r(194) = .70$, $p < .001$, and the average offensiveness rating for all 12 gender stereotypes (Cronbach's $\alpha = .88$).

Stereotype personalizing. Participants were asked to consider the following:

With some stereotypes, you may feel it is personally important to you to work to see them eliminated. For others, you will feel it is less important to see them eliminated. We would like to know how important it is to you to see each of the stereotypes eliminated from society.

They then rated how important it was to them to help eliminate each of the 12 stereotypes rated for importance, on

TABLE 1
Associations between Ratings of the Math and Spatial Ability Stereotypes

Variable	2	3	Mean	SD	Median	MAD
1. Endorsement	-0.27*	-.04	2.52	2.00	2.29	1.45
2. Offense	—	.64*	4.96	2.65	5.01	2.00
3. Personalizing	—	—	4.89	1.83	5.94	1.50

Note. Endorsement, offense, and personalizing were measured on 11-point scales, with high values indicating high degree of endorsement, offense taken, and personal desire to eliminate. *SD*, standard deviation of the mean; *MAD*, median absolute deviation. Asterisk (*) indicates a significant association at $p < .01$.

a unipolar 11-point rating scale ranging from 0 to 10. Underneath the numbers 0, 3, 6, and 9 were centered the labels *not at all important*, *somewhat important*, *very important*, and *extremely important*, respectively. As before, we computed the average ratings for the math and spatial ability stereotypes, $r(194) = .68$, $p < .001$, and we computed the average rating for all 12 stereotypes (Cronbach's alpha = .89).

Results

Descriptive statistics. Table 1 provides a sense of how participants viewed the math stereotype. It shows the mean level of stereotype belief, stereotype offense, and stereotype personalizing, along with the correlation matrix between

these three measures. The sample as a whole did not show a high degree of belief. The mean ($M = 2.52$, $SD = 2.00$) and median (median = 2.29, median absolute deviation [MAD] = 1.45) both indicate an endorsement just short of the value of 3 (= *some truth*). Inspection of the full distribution suggests, however, that many in the sample did show moderate to high levels of endorsement. Figure 2 presents the frequency distribution of the responses to the two stereotype endorsement items. For ease of interpretation, the evaluative anchors for the 11 points of response scales are included along the x axis. Although 19% of the sample rated both the math and spatial stereotypes at the level of 0 (= *not at all true*), 17% gave ratings at a level that was at or above the scale midpoint of 5, and 41% gave ratings at a level that

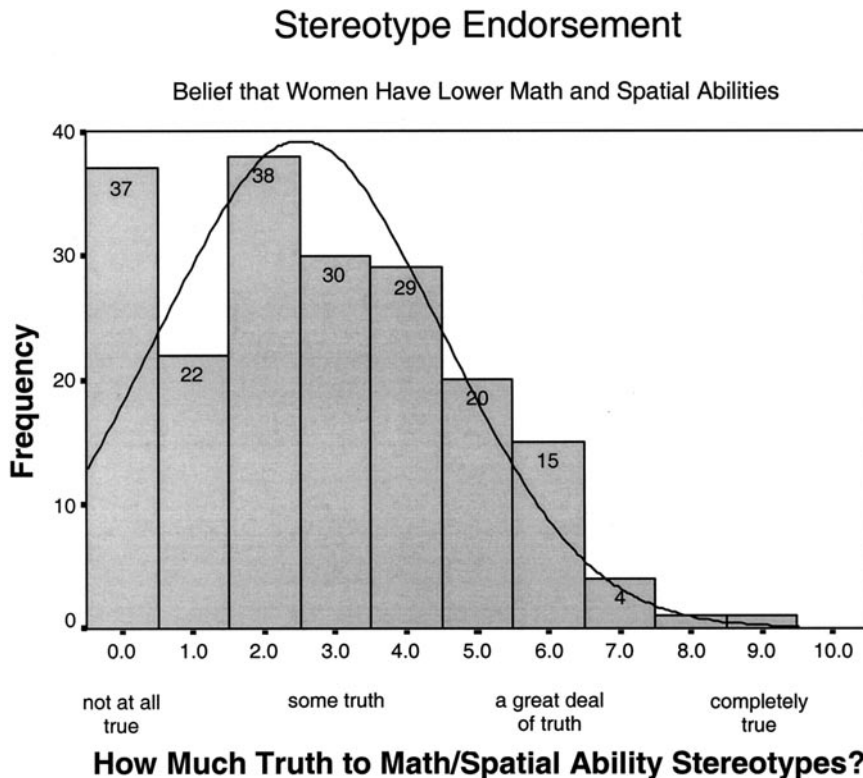


FIG. 2. Distribution of endorsement of the average endorsement of the math and spatial stereotypes.

TABLE 2
Ability Evaluation as a Function of Comparison Information and Stereotype Endorsement

Independent variables	<i>B</i>	<i>t</i>	Significance
Block 1: Main effects			
Constant	7.47	36.11	.00
Comparison information	0.00	0.01	<i>ns</i>
Stereotype endorsement	0.03	0.51	<i>ns</i>
Block 2: Simple main effects and cross-product			
Constant	7.17	30.29	.00
Comparison information	0.67	1.94	.05
Stereotype endorsement	0.15	2.02	.05
Comparison Information × Stereotype Endorsement	-0.27	2.48	.01

Note. Shown are unstandardized coefficients for equation regressing the ability evaluation on comparison information, average endorsement of the math and spatial ability stereotypes, and the multiplicative cross-product of the two. Comparison information is dummy coded such that 0 = women-down condition and 1 = women-up condition. Endorsement is coded such that higher scores indicate a greater belief in the gender stereotype. Ability evaluation can range from 0 to 10, with higher values indicating more positive self-evaluation. All *t* values are assessed with degrees of freedom = 190.

was at or above the level of 3 (= *some truth*). It appears, then, that many in the sample were at least open to the idea that their own gender was at a disadvantage relative to the male gender. Although these ratings do not suggest strong support for the stereotype in the sample as a whole, it is important to remember that these are ratings of endorsement for a negative stereotype about participants' own gender. These ratings are quite high when compared to the traditional view that people are biased to endorse positive self-views (Taylor & Brown, 1988; cf. Jost & Banaji, 1994).

Despite this possibly counterintuitive level of endorsement, the fact remains that the sample as a whole did not have a degree of endorsement that was high in absolute magnitude. This might explain why the results in Study 1 were weaker than anticipated. The level of endorsement in that study might not have been high enough to reveal reference frame effects in analyses conducted on the entire sample. The full distribution of the current sample indicated, however, that many participants showed moderate to high levels of stereotype endorsement. As a result, significant reference frame effects should be pronounced among this subgroup.

In comparison to stereotype endorsement, the level of stereotype offense was high in an absolute sense ($M = 4.96$, $SD = 2.65$; median = 5.01, $MAD = 2.00$), as was the level of stereotype personalizing ($M = 4.89$, $SD = 1.83$; median = 5.94, $MAD = 1.50$). The correlation matrix in Table 1 reveals that stereotype offense and stereotype personalizing were highly correlated, $r(194) = .64$, $p < .01$. This suggests that those who were most offended by the math stereotype were also the ones who most wanted to see it eliminated.² Stereotype endorsement was reasonably inde-

pendent of these two constructs, although it did show a negative correlation with stereotype offense, $r(194) = -.27$, $p < .01$. This indicates that the women who had the highest level endorsement of the math and spatial stereotypes were also the ones who were the least offended by them. Importantly, however, stereotype endorsement showed enough independence from these two alternative moderators that its effects could be disentangled from them.

Stereotype endorsement as moderator. The primary prediction was that endorsement of the stereotype would moderate the effect of the social comparison manipulation on self-evaluation. To test this, ability evaluation was regressed on stereotype endorsement, a dummy code for experimental condition, and the multiplicative cross-product of stereotype endorsement and experimental condition (Jacard, Turrissi & Wan, 1990). For clarity of presentation, the dummy code for experimental condition was constructed to indicate whether or not the two women outperformed the man. Thus, those in the women-up condition were given a value of 1 and those in the women-down condition were given a value of 0. To first test for the main effects of stereotype endorsement and experimental condition, however, these two variables were entered in the first block of a hierarchical regression, with the cross-product entered in the second step.

The unstandardized regression coefficients are shown in Table 2. The first block revealed that there were no main effects of stereotype endorsement or experimental condition, F 's < 1.0. The second block revealed that, consistent with predictions, stereotype endorsement moderated the ef-

² In fact, this high degree of association between stereotype offense and stereotype personalizing was roughly equivalent to the interitem correla-

tion in the two scales. This suggests that they could be viewed as a single construct. As a result, all analyses using these constructs were repeated using the combined average of these two scales. This procedure did not alter the conclusions.

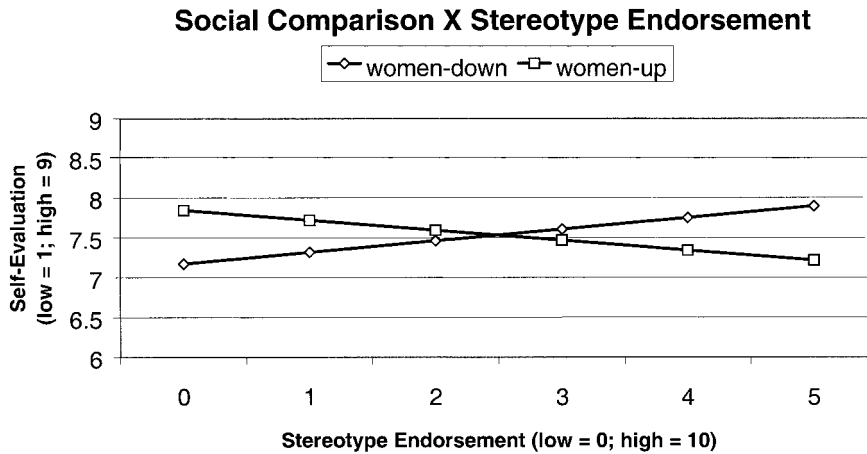


FIG. 3. Ability evaluations in Study 2 as a function of social comparison and stereotype endorsement.

fect of experimental condition on self-evaluation. The direction of the unstandardized regression coefficient for the cross-product indicated that, as stereotype endorsement increased, self-evaluations in the women-up condition diminished relative to those in the women-down condition, $B = -.27$, $t(190) = 2.48$, $p < .01$. More specifically, the magnitude of the unstandardized regression coefficient indicated that the mean rating in the women-up condition dropped an additional 0.27 unit below the mean rating in the women-down condition for every 1 unit increase in stereotype endorsement.³

The nature of the Social Comparison \times Stereotype Endorsement interaction can be further understood by inspecting Fig. 3. This figure plots the estimated ability rating at different levels of stereotype endorsement, broken down by experimental condition. As predicted, stereotype endorsement determined whether or not self-evaluations increased or decreased in the two conditions. When stereotype endorsement was at its lowest point, or equal to 0, ability evaluations were $B = .67$ unit higher in the women-up condition than in the women-down condition. This indicated a significant difference in ability ratings, $t(190) = 1.94$, $p < .05$. Thus, among women who did not believe the stereotype, ability ratings were higher when the two women outperformed a man, even though this meant that participants themselves performed poorly relative to another woman. When stereotype endorsement was at the scale midpoint of 5, however, ability ratings were .68 unit lower in the women-up condition than in the women-down condition. This represented a significant difference, $B = -.68$, $t(190) = 1.96$, $p < .05$, and it indicated that, among women showing a moderate degree of belief in the stereotype,

ability ratings were higher when they performed well relative to another woman. This was true even though the two women performed worse than the man. Interestingly, when the 2-item math ability stereotype scale was replaced with the full 12-item scale, this interaction remained significant, $B = -.35$, $t(193) = 2.43$, $p < .02$. The slight drop in the magnitude of the effect suggests that the specific belief provided the best prediction.

Alternative moderators. To test for the effects of the two alternative moderators, equivalent regression analyses were run replacing stereotype endorsement with stereotype offense and stereotype personalizing. No main effects or interactions were found for these measures. Follow-up analyses also revealed that the previously uncovered moderating effect of stereotype endorsement remained even when these two constructs and the importance of doing well in math were controlled. Assuming valid measurement, these findings argue against the alternative moderators.

Discussion

Study 2 provided clear evidence that endorsement of negative ability stereotypes moderates reactions to social comparison information when these stereotypes have been activated. Female participants who did not believe that men have better math and spatial abilities than women showed higher self-evaluations when they and another woman outperformed a man on a test of math and spatial ability. This occurred even though the other woman had performed better than the participants had performed. This suggests that participants who were low in stereotype endorsement based their self-evaluations on the relative ability of women versus men. As stereotype endorsement increased, however, participants showed higher self-evaluations when they outperformed the other woman on the task. This occurred even though their superior performance occurred in a comparison context in which the male confederate outperformed both of

³ Inspection of residuals revealed one individual who had self-evaluations that were 3 standard deviations from predicted values. Removal of this individual did not alter the significance level of any of the significant interactions reported.

the women. This suggests that participants who were higher in stereotype endorsement based their self-evaluations on their own performance relative to the other woman.

GENERAL DISCUSSION

We began this article by contrasting the social identity view of gender with the reference frame view of gender. The social identity view argues that women will base their self-evaluations on the relative standing of women versus men. The reference frame view argues that women will base their self-evaluations on their personal standing relative to other women and that they will be unconcerned about their own standing relative to men. We hypothesized that negative ability stereotypes, when believed, would cause women to treat their gender as a frame of reference, even in a comparison context that would otherwise encourage them to treat it as a social identity. Our first task was thus to create a comparison context that would cause women to treat their gender as a social identity. This was accomplished by providing female participants with comparison feedback about how their own performances on a test of math and spatial ability compared to the performances of both a male and a female confederate. As predicted, participants had higher ability ratings when they were led to believe that the two women outperformed the man than when it suggested the reverse (see also Sanders et al., 1979). Although this finding was not the main focus of Study 1, it provided important evidence that women will treat their gender as a social identity during comparisons of ability. This adds to a growing experimental literature showing that people often feel enhanced by the superior performance of others when they share a meaningful bond or group identity with them (Blanton et al., 2000; Brown, Novick, Lord, & Richards, 1992; Lockwood & Kunda, 1997; Pelham & Wachsmuth, 1995; for relevant reviews, see Blanton, 2001; Buunk & Ybema, 1997; Tesser, 1988).

As predicted, activation of a negative stereotype altered the contingencies for self-evaluation. When women in Study 1 were reminded of the negative stereotypes that women have worse math and spatial abilities than do men, they showed a tendency to have higher ability evaluations when they performed better than another woman, even if this meant that the two women performed worse than the man. Although this tendency was not statistically significant, Study 2 offered an explanation for this and revealed a group for which the trend was significant. This study showed that belief in the negative math and spatial stereotypes was low for the entire sample but that a sizable minority did show moderate to high belief in the in-group stereotypes. Among this latter group, stereotype activation led to a significant reference frame effect. In other words, they had higher self-evaluations when they performed better

than another woman, even though this meant that the two women performed worse than the man.

It might seem that the activation of gender stereotypes would have an opposite effect and that it would increase the tendency for women to identify with one another during social comparisons. Negative gender stereotypes could increase attention to the intergender implications of comparisons and thereby increase concern for maintaining a positively distinct gender identity. To wit, an effect such as this was predicted and found in our earlier work. Blanton et al. (2000) showed that Black students can identify with one another when they compare with members of their racial in-group in the context of a negative in-group stereotype. Many other researchers have argued for similar effects for women. It has been suggested, for instance, that strong female role models help women to maintain positive self-views and high aspirations in the face of lowered stereotypical expectations for their gender (e.g., Barak, Feldman, & Noy, 1991; Hackett, Esposito, & O'Halloran, 1989; Nauta, Epperson, & Kahn, 1998). Although the current results challenge our own past research and this line of reasoning, they nevertheless follow from a straightforward application of the relevant attributes hypothesis (Goethals & Darley, 1977). This social comparison principle predicts that people will base their self-evaluations on their personal standing relative to in-group standards when group membership is predictive of ability. It follows from this that a woman who endorses a gender stereotype about an ability will use social comparison information to determine whether her own ability is "good for a woman" (Biernat & Manis, 1994).

Generalization. A reasonable concern with this study is that it used a contrived method that might not generalize to a wide array of social comparison contexts. This concern can be broken into two broad categories. First, the effects might not hold if the nature of the social comparison information were altered. Second, the effects might not hold if the nature of the social stereotype were altered. With respect to the first concern, we admit that the feedback method we developed may have limited generality. Because participants were always given conflicting inter- and intragroup social comparison information, we were not able to ascertain how participants might have responded if one part of the comparison compensated for another. We would argue, however, that the current studies were concerned with demonstrating that negative stereotypes can shift an individual to a reference frame view of the in-group *when a social identity view would otherwise predominate*. Based on self-categorization principles, we predicted that our feedback method would cause a female participant to treat her gender as a social identity because it placed her and another woman in either a superior or inferior status relative to a man (Turner, 1985). The effects on ability evaluation supported this prediction. The fact that negative stereotypes lead to reference frame effects using this same comparison method

suggests that the effects of negative stereotypes will generalize to any number of comparison contexts in which individuals would otherwise identify with one another on the basis of a shared social identity.

The second concern is that our effects might not hold if the nature of the social stereotype were altered. We chose to focus on gender and math stereotypes because we believed that they are often endorsed to some degree, even by their targets. Although we predict that our effects will generalize to other in-group stereotypes when they generate similar levels of endorsement, it is possible that endorsement such as this occurs only rarely. Future research will need to address this question, however. It may be that, contrary to intuition, endorsement of in-group stereotypes is often quite high. In fact, a number of theoretical traditions argue that targets of negative stereotypes do internalize the stereotypes about their groups (see Jost & Banaji, 1994). Still others suggest that the pervasive nature of social stereotypes cause individuals to treat them as if they are true, even if they explicitly reject them (e.g., Steele & Aronson, 1995). If future research indicates common acceptance of many in-group stereotypes, we expect that the reference frame shift we have observed will prove to be quite common as well.

Self-esteem versus stereotype maintenance. These studies point to possible tensions between the desire to see the in-group as positively distinct from the out-group and the competing desire to see the self as positively distinct from the in-group. The current study proposed a resolution to this tension by arguing that people refer back to an underlying belief structure (social stereotypes) to determine which aspect of the comparison situation is the most informative guide for self-evaluation. One surprising implication of this resolution is that participants' belief systems directed their social perceptions in ways that appeared to be quite independent of the ultimate effects on self-evaluation. A plausible alternative to the resolution we proposed might have been that individuals would resolve the tension in whatever manner was most likely to enhance or maintain self-esteem. It might seem, for instance, that people presented with competing inter- and intragroup comparison information would focus on whichever piece of information put the self in the most favorable light. Thus, participants in the women-up condition could have focused their attention on the intergroup implications of the social comparison feedback and felt good about the superior performance of the women relative to the man. By contrast, participants in the women-down condition could have focused their attention on the intragroup implications of the social comparison feedback and felt good about their own superior performance relative to the other woman.

This alternative to our predictions is consistent with the theoretical view arguing that people typically bias their social perceptions to maintain self-esteem (e.g., Taylor &

Brown, 1988). Within the social comparison literature, it is best represented in Buunk and Ybema's (1997) identification contrast model. Like other comparison frameworks (Taylor & Lobel, 1989; Tesser, 1988; Wood & Taylor, 1991), the identification contrast model emphasizes personal freedom to use social comparison information to further perceptual goals, particularly the goal of esteem maintenance. Buunk and Ybema (1997) have shown compelling instances in which people maintain self-esteem by selectively focusing their attention on their similarities with upward targets and on their differences with downward targets. As our two studies demonstrated, however, people can at times be "trapped" by the comparison context to interpret social comparison information in ways that run counter to esteem maintenance goals (see also Wood, 1989). More specifically, it seems that reactions to others will, at times, be determined by the relatively inflexible dictates of social stereotypes.

One compelling aspect of our methodology was that the women in the women-up condition were presented with information showing them that women can outperform men on measures of math and spatial ability. If the participants were so motivated, they might have used this information to justify a strategic suspension of belief in the in-group stereotype. If participants in the women-up condition had done this, they could have benefited in at least two ways. First, they would have been released from a stereotypical belief that puts women at a disadvantage relative to men, making it easier for them to maintain positive self-views (cf. Crocker & Major, 1989). Second, they could have focused attention on their association with the other women who had done well, thereby feeling enhanced by the superior performance of the women relative to the man. Instead of suspending belief, however, participants who were reminded of the stereotype reacted to stereotype-inconsistent feedback as if the stereotype were true. The fact that they did this may show just how pervasive negative stereotypes can be. Once individuals internalize in-group stereotypes, they may rely on them to interpret the very situations that would otherwise suggest to them that the stereotypes are false.

The adaptive view on stereotype retention. This analysis might then suggest that the women who endorsed gender stereotypes responded in a way that was not at all adaptive. They ignored comparison information that argued against the stereotype, even in a context in which stereotype retention put them at risk of diminished self-esteem. Although it is tempting to draw this conclusion, one must be careful about applying such logic. To know what is "adaptive" in this particular situation, one must know whether it is better for these women to retain their prior beliefs or to hold positive self-views. At first blush, it might seem straightforward enough; women should simply drop their stereotypes so that they can hold more positive self-views. This reasoning falls apart, however, if the stereotypes the women

are dropping are actually true (see Oakes, Haslam, & Turner, 1994).

We suggest two ways in which the stereotypes can possess "reality," and either one could justify the continued use of the stereotypes, even in the face of contradictory evidence. First, stereotypes can possess some *objective reality*. With respect to math and spatial ability stereotypes, opinion in the psychological community is mixed. Some researchers have argued that there are biological determinants of gender differences in academic performance (e.g., Benbow & Stanley, 1983), whereas others have taken issue with this research (e.g., Bellisari, 1989) or focused greater attention on the social determinants (e.g., Spencer et al., 1999). We could try to argue for one side or the other in this debate. Then, we could use our own opinion to advocate either that women should drop their belief in these stereotypes or that they should not. We do not do this because, independent of any objective reality, these stereotypes can possess a compelling *social reality*.

The study participants who believed the stereotypes about their gender held a belief system that is not entirely uncommon. In fact, some theoretical perspective in psychology would argue that it would not be possible for these women to believe these stereotypes if they had not had some contact in the past with individuals who also held these beliefs (see Hardin & Higgins, 1996). In the case of these particular stereotypes, participants' beliefs probably developed from meaningful social interactions with important reference others (Eccles et al., 1990; Frome & Eccles, 1998; Jacobs & Eccles, 1992). Their stereotypes may therefore serve important social functions among the subset of participants who believe in them, as expressed by tendencies such as the preference for social contact with similarly minded others who share these same beliefs. If so, the response of our participants could be socially adaptive, even if the objective validity of the stereotype is in question.

To illustrate, it is useful to return to the stereotype used to describe the Biernat and Manis (1994) shifting standards model—that women are shorter than men. This particular stereotype is less controversial than the stereotypes about academic abilities, and so most would find it easier to work with the assumption that the stereotype has both an objective and a social reality. If the height stereotype is an accurate statement about the relative stature of men and women, it is arguably adaptive (and justified) for women to restrict their definitions of "tall" and "short" to same-gender norms. It would also seem to be adaptive for them to retain their height stereotype in the face of isolated instances of disconfirming evidence. Take as a specific illustration a fictitious situation inspired by our feedback methodology. It would not seem adaptive for two women encountering a man who is shorter than they are to abandon their well-learned stereotype that women are typically shorter than men. Doing so would run in the face of their accumulated

wisdom about the distribution of height across genders. It would also put their opinion at odds with the opinions of others in their lives who probably endorse the height stereotype. More to the point, if these women live in a society where men truly are taller than women, where height is socially valued, and where height is a common basis for feelings of self-worth, these women should not to allow their one interaction with a short man to initiate a new interest in cross-gender height comparisons. If this were to happen, these women would open themselves up to many future social comparisons in which they would be at a disadvantage relative to men. Instead of abandoning their stereotype, these women would find it easier to maintain positive and socially valid self-views if they restrict their frame of reference to same-gender norms. If they do this, they can take encouragement when they feel "tall for a woman" and be unconcerned when they fail to measure up to men. The downside is that when they encounter tall women, they will feel short.

We point out the distinction between objective and social reality to show that either can justify the use of prior stereotypes to interpret stereotype-inconsistent information, even when doing so causes personal distress. We rush to point out, however, that neither objective nor social reality is sufficient grounds for labeling such responses as "adaptive." One could argue that it is maladaptive, for instance, for individuals to adjust their perceptions to current realities when doing so reinforces the very situations that put them or their groups at a disadvantage (Blanton, George, & Crocker, 2001; see Major, 1994). Although social stereotypes may at times possess compelling realities, this is not to say that these truths may not be open to change. In such instances, it might benefit individuals to look beyond their current arrangements. In the process, they might see how they can help their groups to move to new and more satisfying frames of reference.

REFERENCES

- Abrams, D., Sparkes, K., & Hogg, M. A. (1985). Gender salience and social identity: The impact of sex of siblings on educational and occupational aspirations. *British Journal of Educational Psychology*, *55*, 224–232.
- Abrams, D., Thomas, J., & Hogg, M. A. (1990). Numerical distinctiveness, social identity and gender salience. *British Journal of Social Psychology*, *29*, 87–92.
- Barak, A., Feldman, S., & Noy, A. (1991). Traditionality of children's interests as related to their parents' gender stereotypes and traditionality of occupations. *Sex Roles*, *24*, 511–524.
- Bellisari, A. (1989). Male superiority in mathematical aptitude: An artifact. *Human Organization*, *48*, 273–279.
- Benbow, C. P., & Stanley, J. C. (1983). Sex differences in mathematical reasoning ability: More facts. *Science*, *222*, 1029–1030.
- Biernat, M., & Manis, M. (1994). Shifting standards and stereotype based judgments. *Journal of Personality and Social Psychology*, *66*, 5–20.
- Biernat, M., Manis, M., & Nelson, T. E. (1991). Stereotypes and standards

- of judgment. *Journal of Personality and Social Psychology*, **60**, 485–499.
- Blanton, H. (2001). Evaluating the self in the context of another: The three-selves model of social comparison assimilation and contrast. In G. B. Moskowitz (Ed.), *Cognitive social psychology: The Princeton Symposium on the Legacy and Future of Social Cognition* (pp. 75–87). Mahwah, NJ: Erlbaum.
- Blanton, H., Crocker, J., & Miller, D. T. (2000). The effects of in-group versus out-group social comparison on self-esteem in the context of a negative stereotype. *Journal of Experimental Social Psychology*, **36**, 519–530.
- Blanton, H., George, G., & Crocker, J. K. (2001). Contexts of system justification and system evaluation: Exploring the social comparison strategies of the (not yet) contented female worker. *Group Processes and Intergroup Relations*, **4**, 127–138.
- Brown, J. D., Novick, N. J., Lord, K. A., & Richards, J. M. (1992). When Gulliver travels: Social context, psychological closeness, and self-appraisals. *Journal of Personality and Social Psychology*, **62**, 717–727.
- Buunk, B. P., & Ybema, J. F. (1997). Social comparison and occupational stress: The identification-contrast model. In B. P. Buunk & F. X. Gibbons (Eds.), *Health, coping, and social comparison* (pp. 359–388). Hillsdale, NJ: Erlbaum.
- Crocker, J., & Major, B. (1989). Social stigma and self-esteem: The self-protective properties of stigma. *Psychological Review*, **96**, 608–630.
- Davis, J. A. (1966). The campus as a frog pond: An application of the theory of relative deprivation to career. *American Journal of Sociology*, **72**, 17–31.
- Dixon, J., & Seron, C. (1995). Stratification in the legal profession: Sex, sector, and salary. *Law and Society Review*, **29**, 381–412.
- Eagly, A. H. (1987). *Sex differences in social behavior: A social role interpretation*. Hillsdale, NJ: Erlbaum.
- Eberhardt, J. L., & Fiske, S. T. (1994). Affirmative action in theory and practice: Issues of power, ambiguity, and gender versus race. *Basic and Applied Social Psychology*, **15**, 201–220.
- Eccles, J. S., Jacobs, J. E., & Harold, R. D. (1990). Gender role stereotypes, expectancy effects, and parents' socialization of gender differences. *Journal of Social Issues*, **46**, 183–201.
- Feldman, N. S., & Ruble, D. N. (1981). Social comparison strategies: Dimensions offered and options taken. *Personality and Social Psychology Bulletin*, **7**, 11–16.
- Frieze, I. H., Olson, J. E., & Good, D. C. (1990). Perceived and actual discrimination in the salaries of male and female managers. *Journal of Applied and Social Psychology*, **20**, 46–67.
- Frome, P. M., & Eccles, J. S. (1998). Parents' influence on children's achievement-related perceptions. *Journal of Personality and Social Psychology*, **74**, 435–452.
- Glick, P., & Fiske, S. T. (1996). The ambivalent sexism inventory: Differentiating hostile and benevolent sexism. *Journal of Personality and Social Psychology*, **76**, 491–512.
- Goethals, G. R., & Darley, J. M. (1977). Social comparison theory: An attributional approach. In J. M. Suls & R. L. Miller (Eds.), *Social comparison processes: Theoretical and empirical perspectives* (pp. 259–278). New York: Hemisphere.
- Gonzales, P. J., Blanton, H., & Williams, K. J. (in press). *The effects of stereotype threat and double-minority status on the test performance of Latino women*. *Personality and Social Psychology Bulletin*.
- Hackett, G., Esposito, D., & O'Halloran, M. S. (1989). The relationship of role model influences to the career salience and educational and career plans of college women. *Journal of Vocational Behavior*, **35**, 164–180.
- Hagan, J. (1990). The gender stratification of income inequality among lawyers. *Social Forces*, **68**, 835–855.
- Hardin, C. D., & Higgins, E. T. (1996). Shared reality: How social verification makes the subjective objective. In R. M. Sorrentino & E. T. Higgins (Eds.), *Handbook of motivation and cognition* (Vol. 3, pp. 28–84). New York: Guilford.
- Heatherington, T. F., & Polivy, J. (1991). Development and validation of a scale for measuring state self-esteem. *Journal of Personality and Social Psychology*, **60**, 895–910.
- Inzlicht, M., & Ben-Zeev, T. (2000). A threatening intellectual environment: Why females are susceptible to experiencing problem-solving deficits in the presence of males. *Psychological Science*, **11**, 365–371.
- Jaccard, J., Turrissi, R., & Wan, C. K. (1990). *Interaction effects in multiple regression* (Sage University Paper series on Quantitative Applications in the Social Sciences, No. 07-072). Newbury Park, CA: Sage.
- Jacobs, J. E., & Eccles, J. S. (1992). The impact of mothers' gender-role stereotypic beliefs on mothers' and children's ability perceptions. *Journal of Personality and Social Psychology*, **63**, 932–944.
- Jost, J. T., & Banaji, M. R. (1994). The role of stereotyping in system-justification and the production of false consciousness. *British Journal of Social Psychology*, **33**, 1–27.
- 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. (1994). From social inequality to personal entitlement: The role of social comparisons, legitimacy appraisals, and group membership. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 26, pp. 293–348). San Diego: Academic Press.
- Major, B., & Forcey, B. (1985). Social comparisons and pay evaluations: Preferences for same-sex and same-job image comparisons. *Journal of Experimental Social Psychology*, **21**, 393–405.
- Major, B., Sciacchitano, A., & Crocker, J. (1993). In-group versus out-group comparisons and self-esteem. *Personality and Social Psychology Bulletin*, **19**, 711–721.
- Marsh, H., & Parker, J. (1984). Determinants of students' self-concept: Is it better to be a relatively large fish in a small pond even if you don't learn to swim as well? *Journal of Personality and Social Psychology*, **47**, 213–231.
- McFarland, C., & Buehler, R. (1995). Collective self-esteem as a moderator of the frog-pond effect in reactions to performance feedback. *Journal of Personality and Social Psychology*, **68**, 1055–1070.
- Miller, C. T. (1984). Self-schemas, gender, and social comparison: A clarification of the related attributes hypothesis. *Journal of Personality and Social Psychology*, **46**, 1222–1228.
- Nauta, M. M., Epperson, D. L., & Kahn, J. H. (1998). A multiple-groups analysis of predictors of higher level career aspirations among women in mathematics, science, and engineering majors. *Journal of Counseling Psychology*, **45**, 483–496.
- Oakes, P. J., Haslam, A., & Turner, J. C. (1994). *Stereotyping and social reality*. Oxford, UK: Blackwell.
- Pelham, B. W., & Wachsmuth, J. O. (1995). The waxing and waning of the social self: Assimilation and contrast in social comparison. *Journal of Personality and Social Psychology*, **69**, 825–838.
- Peplau, L. A. (1983). Roles and gender. In H. H. Kelley, E. Berscheid, A. Christensen, J. H. Harvey, T. L. Huston, G. Levinger, E. McClintock, L. A. Peplau, & D. R. Peterson (Eds.), *Close relationships* (pp. 220–264). New York: Freeman.
- Sanders, G. S., Gastorf, J. W., & Mullen, B. (1979). Selectivity in the use of social comparison information. *Personality and Social Psychology Bulletin*, **5**, 377–380.

- Spencer, S. J., Steele, C. M., & Quinn, D. M. (1999). Stereotype threat and women's math performance. *Journal of Experimental Social Psychology*, **35**, 4–28.
- Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology*, **69**, 797–811.
- Suls, J. M., Gaes, G. G., & Gastorf, J. W. (1979). Evaluating a sex-related ability: Comparison with same–opposite and combined-sex norms. *Journal of Research in Personality*, **13**, 294–304.
- Tajfel, H. (1978). *Differentiation between social groups: Studies in the social psychology of intergroup relations*. London: Academic Press.
- Tajfel, H., & Turner, J. C. (1979). An integrative theory of intergroup conflict. In S. Worchel & W. G. Austin (Eds.), *The social psychology of intergroup relations* (pp. 33–47). Pacific Grove, CA: Brooks/Cole.
- Taylor, S. E., & Brown, J. D. (1988). Illusion and well-being: A social psychological perspective on mental health. *Psychological Bulletin*, **103**, 193–210.
- Taylor, S. E., & Lobel, M. (1989). Social comparison activity under threat: Downward evaluation and upward contacts. *Psychological Review*, **96**, 569–575.
- 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.
- Turner, J. C. (1985). Social categorization and the self-concept: A social cognitive theory of group behavior. In E. J. Lawler (Ed.), *Advances in group processes* (Vol. 2, pp. 77–122). Greenwich, CT: JAI.
- Turner, J. C., Hogg, M. A., Oakes, P. J., Reicher, S., & Wetherell, M. S. (1987). *Rediscovering the social group: A self-categorization theory*. Oxford, UK: Basil Blackwell.
- Turner, J. C., Oakes, P. J., Haslam, S., Haslam, A., & McGarty, C. (1994). Self and collective: Cognition and social context. *Personality and Social Psychology Bulletin*, **20**, 454–463.
- Wood, J. V. (1989). Theory and research concerning social comparisons of personal attributes. *Psychological Bulletin*, **106**, 231–248.
- Wood, J. V., & Taylor, K. L. (1991). Serving self-relevant goals through social comparison. In J. Suls & T. A. Wills (Eds.), *Social comparison: Contemporary theory and research* (pp. 23–49). Hillsdale, NJ: Erlbaum.