Sex Roles, Vol. 53, Nos. 9/10, November 2005 (© 2005)
DOI: 10.1007/s11199-005-7735-2

Effects of Salient Multiple Identities on Women's Performance Under Mathematics Stereotype Threat

Dana M. Gresky,[1] Laura L. Ten Eyck,[1] Charles G. Lord,[1,3] and Rusty B. McIntyre[2]

Previous research on affective extremity and social identity complexity suggested that women's mathematics stereotype threat might be alleviated by reminding individual women of their multiple roles and identities, most of which would presumably be unrelated and thus impervious to negative stereotypes regarding math performance. To test this hypothesis, we primed the relevant stereotype and then asked men and women college students to draw self-concept maps with many or few nodes. When they drew no maps or maps with few nodes, highly math-identified women scored significantly worse than highly math-identified men on a subsequent Graduate Record Examination-like math test, but when they drew maps with many nodes, they scored as well as those men. Theoretical and practical implications of the results are discussed.

KEY WORDS: stereotype threat; social identity; self-complexity.

Stereotype threat can be defined as "the immediate situational threat that derives from the broad dissemination of negative stereotypes about one's group—the threat of possibly being judged and treated stereotypically, or of possibly self-fulfilling such a stereotype" (Steele & Aronson, 1995, p. 798). When a situation creates stereotype threat, members of the negatively stereotyped group perform poorly relative to their actual level of competence (Steele, 1997, 1998). Performance deficits from stereotype threat have been found for African Americans on verbal tests (Blascovich, Spencer, Quinn, & Steele, 2001; Steele & Aronson, 1995), Latino/a students on spatial ability tasks (Gonzales, Blanton, & Williams, 2002), people of lower income on verbal tests (Croizet & Claire, 1998), Black students in tasks that require "sports intelligence" (Stone, Lynch, Sjomeling, & Darley, 1999), White students in tasks that require "natural athletic ability" (Stone et al., 1999), and women on mathematics tests (Inzlicht & Ben-Zeev, 2000; Keller & Dauenheimer, 2003; Quinn & Spencer, 2001; Schmader, Johns, & Barquissau, 2004; Spencer, Steele, & Quinn, 1999).

Researchers have discovered several effective techniques to alleviate the performance deficits associated with stereotype threat. These techniques include lessening the importance of the task (Croizet & Claire, 1998; Steele & Aronson, 1995), reducing the salience of the stereotype (Spencer et al., 1999), claiming the test is not susceptible to the stereotype (Walsh, Hickey, & Duffy, 1999), providing excuses for poor performance (Brown & Josephs, 1999; Stone et al., 1999), reducing opportunities to self-handicap (Keller, 2002), altering ability conceptions from static to fluid (Aronson, Fried, & Good, 2002), and presenting group members with information that suggests that members of their group can be successful on the task (Marx & Roman, 2002; McIntyre, Paulson, & Lord, 2003).

The alleviation technique of particular interest for the present research involves making salient other social identities. In a relevant study (Shih, Pittinsky, & Ambady, 1999, Study 1), Asian American women were reminded either of their

[1]Texas Christian University, Fort Worth, Texas.
[2]Amherst College, Amherst, Massachusetts.
[3]To whom correspondence should be addressed at Department of Psychology, Texas Christian University, Box 298920, Fort Worth, Texas 76129; e-mail: c.lord@tcu.edu.

identity as women or of their identity as Asians, and then took a math test. Those who were reminded of their identity as Asians got a larger percentage of items correct than did those who were reminded of their identity as women. Making a more competent group identity salient alleviated the performance deficits that usually occur when individuals are under stereotype threat.

Unfortunately, few targets of stereotype threat have available to them an alternate social identity that has exactly the opposite implications. Non-Asian women, for example, may not have an alternate social identity that has the reputation of being especially talented at math. Would it help the average woman's math performance to think about social identities other than being a woman, even if those other identities suggested no extraordinary math talent? Might women's math performance under stereotype threat benefit from being reminded that they are not only women, but also members of various family, friendship, school, or activity groups? Such a benefit was suggested by research on affective extremity and social identity complexity.

Affective Extremity

The affective extremity hypothesis holds that objects are evaluated more moderately when they are conceptualized along multiple rather than few dimensions. In a relevant study, participants who were directed to think about a law school applicant's essay along six dimensions later evaluated a weak applicant less negatively (and a strong applicant less positively) than did participants who were directed to think about the same essay along only two dimensions (Linville & Jones, 1980, Experiment 4). Linville (1985) subsequently applied the affective extremity hypothesis to evaluations of one's own performance. In one study (Linville, 1985, Experiment 1), participants sorted personality traits into categories that described themselves. Those who formed many categories later reacted less negatively to bogus failure feedback than did those who formed fewer categories. In another study (Linville, 1985, Experiment 2), participants who sorted self-relevant labels (e.g., woman, wife, student, and romantic partner) into many categories subsequently reported more stable daily moods than did those who used fewer categories.

Linville (1987) depicted thinking about oneself along multiple dimensions as a "cognitive buffer" against the negative impact of stressful life events. Individuals who think about many aspects of their self-concepts have a smaller proportion of those aspects affected by any one negative event than do individuals who think about few aspects of their self-concepts, and thus they are relatively insulated from specific stressful events. This buffering effect of thinking about the self along many versus few dimensions has been supported in numerous studies, but only when the event in question is perceived as self-relevant and genuinely stressful, not just a minor hassle (Rafaeli-Mor & Steinberg, 2002).

Although the affective extremity hypothesis has not previously been applied to situations that elicit stereotype threat, its implications seem clear. If women, for instance, could be induced to think about their multiple roles or identities other than as a member of the general category "women," doing so might have a buffering effect during math tests. Stereotype threat exerts an additional pressure or burden on members of the stereotyped group because poor performance will reflect poorly not only on them but also on their group (Steele, 1997). When they enter certain situations, such as a difficult mathematics test, most students experience some stress, but women, as members of a group that stereotypically performs poorly, experience even greater stress because they are additionally threatened by their group's stereotype. That additional burden might be relieved, and performance might be improved, by reminding women that they have many roles and identities, other than the role of "student." Then stress caused by threat to that one identity would have less impact because a smaller proportion of their salient identities would be threatened. Such an outcome would be consistent with other evidence that women who hold a greater number of social roles experience less psychological distress than women who hold fewer social roles (Gove, 1972; Radloff, 1975; Thoits, 1983).

Social Identity Complexity

Research on social identity complexity also suggests that making multiple social identities salient might improve performance under stereotype threat. People have multiple social identities (Brewer, 1991; Stryker & Statham, 1985; Tajfel, 1978). A typical college woman, for example, might identify herself by her sex, race, ethnicity, social class, religion, sorority, class in school, job, athletic team, club membership, family, and other personal relationships. These multiple identities vary in the probability that each will become salient in different contexts (Turner, 1987). At some times, in some contexts, a woman might be most aware of being a woman, and yet at other times and in other contexts, the same woman might be most aware of being a friend, a runner, a sibling, or a spouse.

Roccas and Brewer (2002) reviewed the situational factors that can cause an individual to acknowledge or to ignore differences among that individual's multiple ingroup identities. These situational factors include temporary distinctiveness (McGuire & McGuire, 1988), cognitive load (Conway, Carroll, Pushkar, & Arbuckle, 1996), stress (Wegner & Wenzlaff, 1996), and ingroup threat (Rothgerber, 1997). A threat to an individual's ingroup can temporarily increase the salience of membership in the threatened group, cause that group to dominate the social identity, induce stress, deplete attentional resources, and lead to a simplified representation of the individual's other social identities (Roccas & Brewer, 2002). When that happens, individuals sometimes react to the threat by emphasizing a different, higher status identity (Roccas, 2003).

Multiple social identities may serve as a buffer against aversive effects of threats to the status of a particular group because individuals "who are simultaneously members of multiple groups that differ in their status can shift their locus of identity between their lower status and higher status groups" (Roccas & Brewer, 2002, p. 103). Women who are under stereotype threat, for instance, might reduce their identification with aspects of being a woman that fit the relevant stereotype, yet remain identified with aspects of being a woman that they perceive as unrelated to, or at least not deficient in, math (Pronin, Steele, & Ross, 2004). When individuals have a larger number of salient social identities, they have greater potential for shifting to a different identity that is not threatened. Salient multiple identities reduce the importance of any one social identity for satisfying an individual's need for a positive self-definition (Brewer, 1991). No previous study, however, has been conducted to test whether reminding women of many other social identities might improve performance under stereotype threat.

One caveat in testing this possibility, however, is that previous research suggests that not all women fall prey to mathematics stereotype threat. Women differ in the extent to which they identify with the domain of mathematics (Smith & White, 2001). Women who are relatively low in math identification often score poorly on difficult math tests whether they are placed under stereotype threat or not (Steele, Spencer, & Aronson, 2002). Women who are relatively high in math identification are the ones who perform at less than their normal levels when placed under math stereotype threat, and they are the ones whose performance might be improved by an effective intervention (Spencer et al., 1999). Thus some stereotype threat researchers have limited participants to those high in domain identification (e.g., Brown & Josephs, 1999, Experiment 3; Davies, Spencer, Quinn, & Gerhardstein, 2002), and others have included unselected participants, but found stereotype threat effects only for those who were highly identified with the relevant domain (e.g., Aronson et al., 1999; Cadinu, Maass, Frigerio, Impagliazzo, & Latinotti, 2003). The present study's predicted alleviation of stereotype threat effects, then, might be qualified by an interaction with women's level of math identification.

METHOD

Participants

One hundred and twenty-nine students (94 women and 35 men) participated for course credit. Although ethnicity and age were not collected, the participants were all recruited from the psychology department's courses, which typically include predominantly White (79%) students between 18 and 22 years of age, more than two-thirds of whom are women. Participants' math identification scores had been obtained as part of a large questionnaire packet administered earlier in the semester. The math identification measure (Smith & White, 2001) contained items such as "I get good grades in math" (on scales from 1 = *strongly disagree* to 5 = *strongly agree*), "How much do you enjoy math-related subjects?" (on scales from 1 = *not at all* to 5 = *very much*), and "How important is it to you to be good at math?" (on scales from 1 = *not at all* to 5 = *very much*).

Procedure

The experimenter began by telling participants that they were going to participate in two unrelated tasks. The first task would involve concept mapping as a tool for representing concepts graphically. The second task would involve a new version of the Graduate Record Examination (GRE) in mathematics. To be sure that all participants were aware of and had recently thought about the relevant stereotype, the experimenter mentioned explicitly that "I'm studying the GRE because of the well-known stereotype that men usually outperform women on math tests." Previous manipulations of stereotype threat have ranged from very subtle (e.g., Ambady, Paik, Steele, Owen-Smith, & Mitchell, 2003; Croizet & Claire, 1998; Inzlicht & Ben-Zeev, 2000; Schmader & Johns, 2003, Experiment 3; Steele & Aronson, 1995, Experiment 4; Stricker & Ward, 2004) to very explicit (e.g., Aronson et al., 1999; Keller, 2002; Keller & Dauenheimer, 2003; Leyens, Désert, Croizet, & Darcis, 2000; Major, Spencer, Schmader, Wolfe, & Crocker, 1998; McIntyre et al., 2003; O'Brien & Crandall, 2003; Smith & White, 2002). We chose the most explicit type of manipulation because we were interested more in interventions that might alleviate stereotype threat than in differences among factors that create stereotype threat. The intent was to maximize the probability that women would perform worse than men in a condition that had only stereotype threat and no intervention, thus providing a "threat only" baseline from which effectiveness of the alleviation manipulation could be assessed.

Next, the students were randomly assigned to participate in one of three experimental conditions: self-concept maps with few nodes, self-concept maps with many nodes, or no self-concept maps. In the self-concept maps with few nodes condition (36 women and 15 men), the experimenter explained that concept maps allow one to depict thoughts on a topic in graphical form by drawing nodes and pathways between the nodes that branch out from a central node (Anderson, 1985; Holley & Dansereau, 1984). The experimenter further explained that "the best way to create a concept map is to find the most basic and fundamentally important characteristics of the topic" and "provide only the most essential information on the topic using only a few connections." To illustrate the procedure, the experimenter showed participants a self-concept map that had only a central node labeled "me" and four other nodes ("school," "hobbies," "family," and "friends") that branched out from the central node. The experimenter then told participants to draw their own self-concept maps on a large piece of paper that was blank except for a small circle labeled "me" in the center.[4]

[4]To equate time spent drawing maps, participants in the few nodes condition also drew several other few nodes maps of well-known television characters.

In the *self-concept maps with many nodes* condition (35 women and 12 men), the experimenter also explained concept mapping but noted that "the best way to create a concept map is to create a complete description of the topic" and "provide plenty of information on the topic using multiple connections." To illustrate the procedure, the experimenter showed participants a self-concept map that had the same five nodes as in the few nodes condition but also had 46 additional nodes that branched out from the central node in a hierarchical network of connections. Then these participants drew self-concept maps with the same materials as in the few nodes condition.

In the *no maps* condition (23 women and 8 men), participants drew no self-concept maps, but instead proceeded directly to the math test. Participants in all three conditions were told that they were "about to complete a portion of the quantitative GRE, and it is diagnostic of your mathematical reasoning abilities. Please carefully complete each of the test items, making sure not to skip any items." The 20-min test consisted of 30 difficult items drawn from sample GRE tests. Difficult items were used because previous research has shown that stereotype threat affects women's performance on relatively difficult, rather than relatively easy, mathematics problems (Spencer et al., 1999). In a thorough debriefing, no participant guessed the experimental hypothesis.

RESULTS

The central conceptual hypothesis was that in situations where women find themselves under mathematics stereotype threat, their performance deficits can be alleviated by reminding them, through their own self-concept maps with many nodes, that they fill many social roles and have many social identities. To test this central hypothesis, it was necessary to show that the self-concept mapping manipulation was sufficient to make salient to women their many roles and identities. The relevant manipulation check concerned whether women in the many nodes condition actually drew self-concept maps with a significantly greater number of nodes than did women in the few nodes condition. Only then could we test whether women who were instructed to draw self-concept maps with many nodes performed better on the math test than did women who were instructed to draw either self-concept maps with few nodes or no self-concept maps.

Manipulation Checks

As in previous research (Aronson et al., 1999, Experiment 2; Cadinu et al., 2003), scores on the math identification measure were divided at their median. On the basis of this median split, 66 participants were classified as relatively low in math identification and 63 participants were classified as relatively high in math identification.[5] A 2 (map condition: few nodes, many nodes) × 2 (gender) × 2 (math identification: low, high) analysis of variance (ANOVA) of the number of nodes in participants' self-concept maps yielded a significant main effect of mapping condition. Participants who were instructed to draw self-concept maps with many nodes included a greater number of nodes in their maps (M = 37.98, SD = 10.59) than did participants who were instructed to draw self-concept maps with few nodes (M = 7.26, SD = 4.65), $F(1, 90) = 243.68$, $p < .001$, $d = 1.32$. The only other significant effect in the ANOVA was an interaction between mapping condition and gender, $F(1, 90) = 7.36$, $p = .008$. The mapping manipulation proved more effective for women than for men. Men's mean values were 32.69 (SD = 9.14) versus 9.09 (SD = 5.70), $F(1, 90) = 57.82$, $p < .001$, $d = 1.35$. Women's mean values were 40.40 (SD = 10.43) versus 6.94 (SD = 4.21), $F(1, 90) = 321.19$, $p < .001$, $d = 1.93$. The important point for testing the central hypothesis was that the manipulation had a significant effect on the number of nodes that women included in their self-concept maps. It remained to assess whether the mapping manipulation had a significant effect on women's mathematics performance.

Mathematics Performance

Previous research on women's math performance under stereotype threat had used three types of performance measures: number correct (e.g., Aronson et al., 1999; Brown & Josephs, 1999; Ford, Ferguson, Brooks, & Hagadone, 2004; Keller & Dauenheimer, 2003; Marx & Roman, 2002; Schmader, 2002; Walsh et al., 1999), percentage of attempted items correct (e.g., Ambady et al., 2003; Davies et al., 2002; Inzlicht & Ben-Zeev, 2003; Keller, 2002; Schmader, 2002; Schmader & Johns, 2003; Shih et al., 1999), and adjusted scores that correct for guessing (e.g., Brown & Pinel, 2003; McIntyre et al., 2003; Quinn & Spencer, 2001; Spencer et al., 1999). Tables I–III display the mean values for these three types of dependent measures in the present experiment. As will be seen, the three types of performance measures suggest similar conclusions.

Table I. Mean Number of Items Correct on a Difficult Math Test by Men and Women who had High or Low Math Identification and Drew No Self-Concept Maps, Self-Concept Maps with Few Nodes, or Self-Concept Maps with Many Nodes

	No maps	Few nodes	Many nodes
Men			
Low math ID	4.00^a	5.00^a (1.87)	4.67^a (1.75)
	$N = 1$	$N = 5$	$N = 6$
High math ID	6.86^b (3.72)	8.33^b (3.83)	7.40^b (2.76)
	$N = 7$	$N = 10$	$N = 6$
Women			
Low math ID	4.72^c (2.22)	4.21^c (3.08)	5.06^c (2.38)
	$N = 18$	$N = 19$	$N = 17$
High math ID	3.80^d (2.59)	4.88^d (2.55)	7.28^e (2.72)
	$N = 5$	$N = 17$	$N = 18$

Note. Standard deviations are in parentheses. The row mean values with different superscripts differed significantly by Tukey's test ($p < .05$).

Number Correct

Table I shows the mean values from a 3 (map condition: few nodes, many nodes, no maps) × 2 (gender) × 2 (math identification: low, high) ANOVA of the number of items correct. The only significant effect in the ANOVA was a main effect of math identification. Participants high in math identification got more items correct ($M = 6.43$, $SD = 3.11$) than did participants low in math identification (M = 4.67, $SD = 2.43$), $F(1, 117) = 7.18$, $p = .008$, $d = .60$. The map condition × gender interaction was not significant, $F(2, 117) = 1{,}634$, $p = .20$.

Previous researchers of stereotype threat effects, however, have reported either marginally significant or no significant interactions between race or gender and the experimental manipulation but have examined the research questions of interest in greater analytic detail (e.g., Blascovich et al., 2001; Brown & Pinel, 2003; Cadinu et al., 2003; Inzlicht & Ben-Zeev, 2003, Experiment 2; Keller, 2002; Keller & Dauenheimer, 2003; McIntyre et al., 2003; Smith & White, 2002, Experiment 1; Steele & Aronson, 1995,

[5]Regression analyses, which treated math identification scores as a continuous variable and the other two predictors (map condition and gender) as dummy variables, yielded the same pattern of significant results as reported for the ANOVAs on both the manipulation check and performance measures.

Experiment 2).[6] Following these precedents, we analyzed men's and women's number correct in separate map condition × math identification ANOVAs followed by simple effects tests for each row of mean values in Table I. For men, as shown on the top two rows, there was only a main effect of math identification. Men who were high in math identification scored better ($M = 7.48$, $SD = 3.29$) than did men who were low in math identification ($M = 4.75$, $SD = 1.66$), $F(1, 117) = 4.36$, $p = .041$, $d = .89$. For women, in contrast, as shown on the two bottom rows of Table I, there was only a significant effect of map condition, $F(2, 117) = 4.22$, $p = .017$. By Tukey's test ($p < .05$), women who were induced to draw self-concept maps with many nodes were correct on more items ($M = 6.20$, $SD = 2.76$) than were women who were induced to draw either self-concept maps with few nodes ($M = 4.53$, $SD = 2.82$) or no self-concept maps ($M = 4.52$, $SD = 2.27$). The latter two conditions did not differ.

For purposes of comparison with the analyses that will be presented for percent correct and adjusted scores, the superscripts on the rows of Table I show the results of simple effects tests for each row of mean values. As can be seen, only the bottom row of the table yielded significant differences among map conditions, $F(2, 117) = 5.04$, $p = .008$. By Tukey's test ($p < .05$), women high in math identification got more items correct when they were induced to draw self-concept maps with many nodes than to draw either self-concept maps with few nodes or no self-concept maps, and the latter two conditions did not differ. For these women, it was not just drawing self-concept maps that improved performance, because drawing self-concept maps with few nodes made no difference. By the number correct measure, high math identification women's performance improved only when they drew self-concept maps that included many social roles and identities.[7]

Table II. Mean Percentage Correct on a Difficult Math Test by Men and Women who had High or Low Math Identification and Drew No Self-Concept Maps, Self-Concept Maps with Few Nodes, or Self-Concept Maps with Many Nodes

	No maps	Few nodes	Many nodes
Men			
Low math ID	17.39^a	22.54^a (11.81)	22.78^a (10.36)
	$N = 1$	$N = 5$	$N = 6$
High math ID	31.37^b (19.43)	33.78^b (18.53)	29.87^b (11.43)
	$N = 7$	$N = 10$	$N = 6$
Women			
Low math ID	20.32^c (10.31)	17.46^c (10.37)	20.38^c (11.64)
	$N = 18$	$N = 19$	$N = 17$
High math ID	16.54^d (10.48)	23.83^d (11.62)	37.14^e (16.47)
	$N = 5$	$N = 17$	$N = 18$

Note. Standard deviations are in parentheses. The row mean values with different superscripts differed significantly by Tukey's test ($p < .05$).

Percent Correct

Table II shows the mean values from a 3 (map condition) × 2 (gender) × 2 (math identification) ANOVA of the percent of attempted items correct. The only significant effect in the ANOVA was a main effect of math identification. Participants high in math identification got a larger percentage of attempted items correct ($M = 31.51$, $SD = 16.05$) than did participants low in math identification ($M = 19.93$, $SD = 10.56$), $F(1, 117) = 11.39$, $p = .001$, $d = .79$. The map condition × gender interaction was not significant, $F(2, 117) = 2.24$, $p =. 11$.

As with number correct, we analyzed men's and women's percent correct scores in separate map condition × math identification ANOVAs followed by simple effects tests for each row of mean values in Table II. For men, as shown on the top two rows, there was only a main effect of math identification. Men who were high in math identification scored better ($M = 36.04$, $SD = 15.95$) than did men who were low in math identification ($M = 22.65$, $SD = 10.12$), $F(1, 117) = 7.18$, $p < .01$, $d = .91$. For women, as shown on the two bottom rows, there was a significant map condition × math identification interaction, $F(2, 117) = 3.68$, $p = .028$. Map condition had no

[6]In addition, as shown in Table I, one cell in the three-way ANOVA had only one participant and thus no measurable variance, which raises doubts about the reliability of interaction terms from a three-factor analysis.

[7]Focused comparisons produced results consistent with previous demonstrations that explicit mention of the stereotype induces performance deficits (e.g., Aronson et al., 1999; Keller, 2002; Keller & Dauenheimer, 2003; Leyens et al., 2000; Major et al., 1998; McIntyre et al., 2003; O'Brien & Crandall, 2003; Smith & White, 2002). In the no maps control condition, which was intended to include a threat but no alleviation, women were correct on marginally fewer items ($M = 4.52$, $SD = 2.27$) than men were ($M = 6.50$, $SD = 3.59$), $F(1, 117) = 3.18$, $p = .077$, $d = .68$. As Table I suggests, the difference was marginally significant between women and men high in math identification, $F(1, 117) = 3.73$, $p = .056$, $d = 1.05$, but not between women and men low in math identification, $F(1, 117) = .07$, $p = .192$, $d = .25$. In the few nodes condition, women ($M = 4.53$, $SD = 2.82$) scored significantly worse than men did ($M = 6.82$, $SD = 3.43$), $F(1, 117) = 6.04$, $p = .015$, $d = .79$. In the many nodes condition, women ($M = 6.20$, $SD = 2.76$) scored as well as men did ($M = 6.38$, $SD = 2.73$), $F(1, 117) = .05$, $p = .823$, $d = .06$.

effect on the scores of women low in math identification, $F(2, 117) = .31$, NS. In contrast, as indicated by superscripts on the bottom row of Table II, map condition had a significant effect on the scores of women high in math identification, $F(2, 117) = 7.27$, $p = .001$. By Tukey's test ($p < .05$), women high in math identification scored a higher percent correct when they were induced to draw self-concept maps with many nodes than when they were induced to draw self-concept maps with few nodes or no self-concept maps, which did not differ. Analyses of percent correct proved similar to those for number correct.[8]

Adjusted Scores

We also calculated each participant's adjusted score by subtracting from the number correct one-quarter of the number of attempted items with a wrong answer. Table III shows the mean values from a 3 (map condition) × 2 (gender) × 2 (math identification) ANOVA of these adjusted scores. The only significant effect in the ANOVA was a main effect of math identification. Participants high in math identification got higher adjusted scores ($M = 2.79$, $SD = 3.92$) than did participants low in math identification ($M = -.06$, $SD = 2.79$), $F(1, 117) = 11.62$, $p < .001$, $d = .78$. The map condition × gender interaction was not significant, $F(2, 117) = 1.57$, $p = .21$.

As with number correct and percent correct, we analyzed men's and women's scores in separate map condition × math identification ANOVAs followed by simple effects tests for each row of mean values in Table III. For men, as shown on the top two rows, there was only a main effect of math identification. Men who were high in math identification scored better ($M = 4.05$, $SD = 3.99$) than did men who were low in math identification ($M = .27$, $SD = 2.92$), $F(1, 117) = 8.19$, $p < .01$, $d = 1.03$. For women, as shown on the two bottom rows, there was a significant map condition × math identification interaction, $F(2, 117) = 3.38$, $p = .038$. Map condition had no effect on the scores of women low in math identification, $F(2, 117) = .06$, NS. In contrast, as indicated by superscripts on the bottom row of Table III, map condition had a significant effect on the adjusted scores of women high in math identification, $F(2, 117) = 6.17$, $p = .003$. By Tukey's test ($p < .05$), women high in math identification got higher adjusted scores when they were induced to draw self-concept maps with many nodes than when they were induced to draw self-concept maps with few nodes or no self-concept maps, which did not differ.[9]

Table III. Mean Adjusted Scores on a Difficult Math Test by Men and Women who had High or Low Math Identification and Drew No Self-Concept Maps, Self-Concept Maps with Few Nodes, or Self-Concept Maps with Many Nodes

	No maps	Few nodes	Many nodes
Men			
Low math ID	$-.75^a$ $N = 1$	$.55^a$ (3.17) $N = 5$	$.21^a$ (3.23) $N = 6$
High math ID	3.29^b (4.89) $N = 7$	5.42^b (4.82) $N = 10$	3.78^b (2.91) $N = 6$
Women			
Low math ID	$-.04^c$ (2.71) $N = 18$	$-.36^c$ (3.02) $N = 19$	$-.01^c$ (2.74) $N = 17$
High math ID	-1.05^d (3.09) $N = 5$	$.99^d$ (2.97) $N = 17$	3.94^e (3.71) $N = 18$

Note. Standard deviations are in parentheses. The row mean values with different superscripts differed significantly by Tukey's test ($p < .05$).

Additional Control Condition

These analyses of number correct, percent correct, and adjusted scores, coupled with examination of the pattern of mean values in Tables I–III, suggested that explicit stereotype threat impaired women's performance and did so especially for women high in math identification—a finding

[8]By focused comparisons, in the no maps control condition, which was intended to include a threat but no alleviation, women scored a significantly lower percent correct ($M = 19.50$, $SD = 10.23$) than men did ($M = 31.37$, $SD = 19.43$), $F(1, 117) = 5.03$, $p = .027$, $d = .81$. As Table II shows, the difference was significant between women and men high in math identification, $F(1, 117) = 4.96$, $p = .028$, $d = 1.15$, but not between women and men low in math identification, $F(1, 117) = .05$, $p = .823$, $d = .20$. In the few nodes condition, women ($M = 20.47$, $SD = 11.29$) scored significantly worse than men did ($M = 33.78$, $SD = 18.53$), $F(1, 117) = 8.97$, $p = .003$, $d = .91$, and in the many nodes condition, women ($M = 29.00$, $SD = 16.48$) scored as well as men did ($M = 29.87$, $SD = 11.43$), $F(1, 117) = .05$, $p = .823$, $d = .06$.

[9]By focused comparisons, in the no maps control condition, which was intended to include a threat but no alleviation, women had significantly lower adjusted scores ($M = -.26$, $SD = 2.76$) than men did ($M = 2.78$, $SD = 4.74$), $F(1, 117) = 5.14$, $p = .025$, $d = .83$. As Table III shows, the difference was significant between women and men high in math identification, $F(1, 117) = 5.13$, $p = .025$, $d = 1.18$, but not between women and men low in math identification, $F(1, 117) = .04$, $p = .842$, $d = .22$. In the few nodes condition, women ($M = .28$, $SD = 3.03$) scored significantly worse than men did ($M = 3.20$, $SD = 4.70$), $F(1, 117) = 6.75$, $p = .011$, $d = .80$, and in the many nodes condition, women ($M = 2.03$, $SD = 3.80$) scored as well as men did ($M = 2.94$, $SD = 3.43$), $F(1, 117) = .17$, $p = .681$, $d = .11$.

similar to results from previous studies (see Steele et al., 2002). Drawing self-concept maps with few nodes proved ineffective in alleviating these performance deficits. Drawing self-concept maps with many nodes, in contrast, allowed women who had been placed under stereotype threat to perform significantly better and as well as men did.

To test an alternative explanation that women's performance improvement was caused by thinking about multiple aspects of any topic, rather than specifically by thinking about multiple self-identities, an additional group of 29 women participated in a separate condition in which they were shown a 51-node map about tools and were then asked to draw a map about "food" that had many nodes. These women, whose math identification scores were not available, drew food maps with a mean value of 31.31 ($SD = 11.55$) nodes. On the subsequent math test, they answered a mean value of 3.52 ($SD = 2.18$) items correct, answered a mean value 20.81% ($SD = 7.13$) of attempted items correct, and had adjusted scores of −.26 ($SD = 2.10$). In one-way ANOVAs that included all women in all four conditions, map condition made a significant difference for number correct, $F(3, 119) = 6.92$, $p < .001$, percent correct, $F(3, 119) = 4.30$, $p = .005$, and adjusted scores, $F(3, 119) = 4.34$, $p < .01$. By Tukey's test ($p < .05$) on all the three performance measures, women who were asked to draw self-concept maps with many nodes scored significantly better than did women in the other three conditions, which did not differ. It was not simply thinking about many aspects of any topic that helped women to overcome performance deficits induced by stereotype threat. In the present experiment, only thinking about many social roles or identities provided math-identified women with an effective alleviation technique.

Number of Items Attempted

The ameliorative effect of self-concept maps with many nodes on women's math test performance might have occurred because making multiple social roles and identities salient altered test-taking strategies. Women who drew self-concept maps with many nodes might have become more cautious and attempted only items on which they were relatively confident or they might have become more bold and attempted items on which they thought that they could exclude at least one of the multiple choice alternatives. This hypothesis about possible changes in test-taking strategy was called into question by the consistent pattern of results across the three performance measures, some of which were corrected for guessing. In addition, in a 2 (map condition) × 2 (gender) × 2 (math identification) ANOVA of the number of items that participants attempted, the only effect of even marginal significance was that participants high in math identification attempted slightly fewer items ($M = 20.98$, $SD = 4.02$) than did participants low in math identification ($M = 23.59$, $SD = 5.42$), $F(1, 117) = 2.92$, $p = .09$, $d = .53$. All other effects and interactions had Fs < 1 and ps $> .49$.

Types of Map Nodes

Finally, one might wonder whether the ameliorative effects of drawing self-concept maps with many nodes depended on which types of nodes the women drew. Recall that participants in both of the two self-concept map conditions were shown sample maps that had a central node for "me" and four main branches labeled "school," "hobbies," "family," and "friends." Nodes in the self-concept maps that participants drew in these two conditions were coded by two of the authors (with 100% agreement) into five node types: academic, activities, family, friends, and "other." The percentage of each node type in each participant's self-concept map was used as the dependent measure in a 2 (map condition) × 2 (gender) × 2 (math identification) × 5 (node type) ANOVA, with repeated measures on the node type factor.

The ANOVA yielded only two significant effects. Participants used some types of nodes more frequently than others, $F(4, 360) = 12.8$, $p < .001$. Mean percentages were 14.47 ($SD = 11.09$) for academic, 25.30 ($SD = 16.31$) for activities, 24.99 ($SD = 16.20$) for friends, 26.95 ($SD = 15.37$) for family, and 8.29 ($SD = 20.76$) for other. The main effect of node type was qualified, however, by a significant map condition × node type interaction, $F(4, 360) = 7.30$, $p < .001$. By simple effects tests, participants in the many nodes condition included a larger percentage of family nodes ($M = 33.63$, $SD = 16.35$) in their self-concept maps than did participants in the few nodes condition ($M = 19.69$, $SD = 10.22$), $F(1, 360) = 15.24$, $p < .001$, $d = .67$. They also included a smaller percentage of "other" nodes ($M = 3.39$, $SD = 13.06$ vs. $M = 13.60$, $SD = 25.85$), $F(1, 360) = 8.18$, $p = .005$, $d = .49$. When percentages of family and other nodes were used as covariates in analyses of number correct, percent correct, and adjusted scores, however, women who drew self-concept maps with many nodes still performed significantly better on all the three performance measures than did women who drew self-concept maps with few nodes. Furthermore, within experimental conditions, there were no significant correlations between women's math test performance and the number of school, activity, friend, or family nodes in their self-concept maps. Differences in node type, therefore, could not explain the effects on women's math performance of making many social roles and identities salient.

Discussion

To summarize, all participants were reminded of the stereotype that women do worse than men on the quantitative GRE test. This explicit manipulation was designed to induce stereotype threat for women, especially for women who identified with and cared about the domain of mathematics (Smith & White, 2001; Steele et al., 2002). After the stereotype had been induced, and before taking the difficult math test, participants were instructed to draw self-concept maps with few nodes, self-concept maps with many nodes, or no self-concept maps. Then all the participants took the math test. Consistent with previous research on stereotype threat (see Steele et al., 2002 for review), women high in math identification, when placed under explicit stereotype threat, scored worse than men high in math identification, after drawing either self-concept maps with few nodes or no maps. As predicted from previous research on affective extremity (Linville, 1985; Linville & Jones, 1980; Rafaeli-Mor & Steinberg, 2002) and self-identity complexity (Brewer, 1991; Roccas & Brewer, 2002), highly math identified women who drew self-concept maps with many nodes performed significantly better, and as well as men did. These results, although subject to several important limitations, suggest the need for further examination of the precise underlying mechanisms.

Limitations

The present research examined not only the central hypotheses but also possible experimental artifacts. The few nodes' condition was included, for instance, to assess whether simply thinking about the self, even in a minimal way, might alleviate women's performance deficits. Drawing self-concept maps with few nodes, however, did not help highly math-identified women to perform better. Similarly, the food maps condition was included to assess whether thinking about many aspects of any topic might alleviate women's performance deficits. Drawing food maps with many nodes, however, did not help. The only effective intervention involved instructing women to draw self-concept maps that included many social roles and identities. No one study, however, could include all the factors that might alter interpretation of the present experiment's results.

No Threat Control

The present experiment, for instance, lacked a no threat control condition, in which participants merely took the math test, with no preliminary instructions. Without a no threat control condition, one cannot be certain that the explicit stereotype threat induction actually produced performance deficits among women high in math identification. One could point to previous demonstrations that even subtle inductions of stereotype threat are sufficient to impair performance for members of negatively stereotyped groups (Steele et al., 2002) and that merely being presented with a difficult math test is sufficient to impair performance for women high in math identification (Spencer et al., 1999), but for complete confidence in the present results, one would need results on exactly the same math test for women who merely took the test without stereotype threat instructions, as would be expected in a normal administration of the GRE.

Math Ability

In addition, the present experiment included no measures of participants' math abilities prior to the experimental situation. Several previous studies of women's mathematics stereotype threat have included measures such as SAT scores, grades in math and math-related courses, and number of math courses completed (e.g., Brown & Josephs, 1999; Inzlicht & Ben-Zeev, 2000; Keller, 2002; Keller & Dauenheimer, 2003; Marx & Roman, 2002; McIntyre et al., 2003). Scores on such measures can be used to establish each participant's baseline math ability, so that performance impairment from the induction of stereotype threat and performance improvement from the experimental intervention can be assessed relative to that baseline. Inclusion of such control variables typically reduces the within-cell variance, increases the sensitivity of statistical analyses, and aids interpretation by identifying the direction of significant results. Presumably, women in the present experiment who reported doing well in math courses (as part of the measure of math identification) actually had higher math ability than women who reported doing poorly in math courses. In fact, the scale's validity in this regard has been well documented (Smith & White, 2001). Even so, it would be desirable to include additional measures of actual math ability in follow-up research.

Small Number

Interpretation of the present results is also hindered by the small number of male participants. The proportion of men and women in the experiment was representative of the sample from which participants were drawn (psychology courses), men were not predicted to differ across conditions in any conceptually meaningful way, and the number of women participants seemed reasonable for testing the central hypotheses about alleviating women's performance deficits, but comparisons of men and women within each mapping condition were not as reliable as one would prefer. Confidence that women (at least women high in math identification) actually performed worse than men in the no maps or threat-only control condition comes at least as much from the fact that our data are consistent with the results of previous research (see Steele et al., 2002, for review) as from our analyses that involved unequal cell sizes for men and women.

Generalizability

Finally, the present experiment was limited in its potential generalizability to other mapping instructions and other stereotyped groups and domains. Participants in the present experiment were shown example maps that involved the self-concept and that (in both the few nodes and many nodes conditions) "started" with the same four nodes: family, friends, hobbies, and school. Without additional studies, it is impossible to know whether the present experiment's results would replicate had the experimenter instead used "food" maps with many versus few nodes as the examples, but then required participants to draw their own self-concept maps. Under these altered conditions, would participants still have drawn significantly different numbers of nodes in their self-concept maps? Would they still have confined their own map nodes primarily to these four types? Most important, would the highly math-identified women still have improved their performance and scored as well as men? Only a follow-up study with fewer or different concept map examples could answer these questions.

Similarly, it would be desirable to examine whether this particular alleviation technique proves effective for other negatively stereotyped groups in other important domains. Without the appropriate follow-up studies, we cannot be sure that drawing self-concept maps with many nodes (or in some other way being reminded of multiple social roles and identities) effectively eliminates performance deficits suffered by such other groups as African Americans and Latinos/as who are placed under stereotype threat regarding various intellectual tasks (Gonzales et al., 2002; Steele & Aronson, 1995), and European American participants who are reminded that they do not score as well as Asians on math tests (Aronson et al., 1999). Presumably all such groups would benefit from being reminded of their many social roles and identities when they find themselves in a situation that involves stereotype threat. There is nothing particularly relevant to women or to mathematics about drawing maps that invoke multiple family relationships, but it remains a possibility, based only on the present experiment's procedure and results, that this particular intervention might elicit underlying mechanisms different from those relevant to other types of stereotype threat.

Underlying Mechanisms

This discussion of underlying mechanisms highlights the one area in which the present results mostly suggest the need for future research. The present experiment and its central predictions were suggested by previous theories of affective extremity and social identity complexity. In addition to these mechanisms, however, one must consider the possible impact of perceived social support, the types of self-identities made salient, changes in perceived self-efficacy, and some combination of these and other underlying mechanisms.

Affective Extremity

Previous research on the affective extremity hypothesis had shown that evaluations are less extreme when people conceptualize a topic along many rather than few dimensions (Linville, 1985; Linville

& Jones, 1980). Conceptualizing a topic along many dimensions provides a cognitive buffer, because shortcomings on a single dimension affect a relatively small proportion of the relevant dimensions. The manipulations in previous studies of affective extremity, however, emphasized dimensions rather than identities. Participants in the present experiment were not explicitly asked to arrange their map nodes into categories or dimensions. It remains an open question, therefore, whether a sorting task similar to that used by Linville (1985) to test the affective extremity hypothesis would also improve performance under stereotype threat. This difference in procedures is relevant to understand the underlying mechanism. Linville and Carlston (1994) distinguished between a cognitive representation's size and its differentiation. With concept maps, it is difficult to measure these two aspects of the representation independently. Very few participants in the present experiment included connecting pathways between one cluster of nodes and another. Future researchers, however, might hold the number of self-concept map nodes constant, but instruct participants to make few versus many cross-cluster connections. Doing so would allow a test of whether the present technique depends more on the sheer number of identities made salient or on associating social identities in a complex, inter-connected cognitive network. In addition, to test the affective extremity mechanism more directly, future researchers might provide bogus feedback about failure on an ambiguous math test and measure the impact of such failure feedback on the self-esteem of women who had been placed under stereotype threat and then asked to draw self-concept maps with few versus many nodes (Gove, 1972; Radloff, 1975; Thoits, 1983).

Social Identity Complexity

Various situational factors, including stress, cognitive load, distinctiveness, and ingroup threat, can temporarily induce relatively simplified self-concepts (Roccas & Brewer, 2002). Situations that elicit stereotype threat might make salient one specific identity (e.g., being a woman in a math test context), and it is exactly the wrong identity in terms of performance expectations (Pronin et al., 2004; Schmader, 2002). Drawing self-concept maps with many nodes might improve performance by reducing the salience or importance of that one identity. In future studies, one could train participants to label their self-concept map nodes with valence, importance, or salience ratings, which would allow a closer examination of reduced distinctiveness of the relevant identity as an underlying mechanism (McGuire & McGuire, 1988).

Such a procedure could also test the idea that one effective defense against stereotype threat involves shifting from the situationally relevant, threatened identity to a higher status alternate identity that is more positively evaluated for the specific task (Roccas, 2003). Not all women have available to them an alternate identity that is known for exceptional math ability (Shih et al., 1999), but all women have available to them a wide variety of alternate identities that are at least not commonly believed to be deficient in mathematics.

Perceived Social Support

Instructions to complete self-concept maps through examples that included "family" and "friends" may have reminded women of their social support networks. Stereotype threat creates stress (Blascovich et al., 2001). Perceived social support lowers perceived stress in threatening situations, reduces the severity of stress-related illness, and speeds recovery (Sarason, Sarason, & Gurung, 1997; Seeman, 1996), whether the perceived support is tangible, informational, or emotional (Cohen, 1988; Lazarus & Folkman, 1984). Perceived social support also affects college students' academic self-worth (Crocker, Luhtanen, Cooper, & Bouvrette, 2003), their ability to remain optimistic when coping with stressful events (Brissette, Scheier, & Carver, 2002), and their resistance to depression (Mickelson, 2001). In addition, perceived social support need not be actually received, but can be beneficial even when it is only implicit and potential (Taylor et al., 2004). For feminine women especially (Reevy & Maslach, 2001), being reminded of the emotional support they have available from friends and family may in itself provide a sense of security when they are placed under stereotype threat and must cope with a difficult math test. It seems possible that the success of the present experiment's alleviation technique might have depended on using "friends" and "family" in the example maps, thus directing attention toward these sources of implicit social support.

Types of Identities Made Salient

It might also prove useful in future research to examine more systematically whether different types of social identities might be responsible for improving performance under stereotype threat. The four types of nodes (school, activities, friends, and family) in the present experiment did not differ in their effects on performance under stereotype threat. Sheer number of nodes was all that mattered, as though each additional social identity either decreased the importance of potential failure or increased the probability of activating at least one positive identity. Future researchers might, however, vary the types of identities that participants are instructed to include in their self-concept maps: identities that involve intersection, dominance, compartmentalization, and merger (Roccas & Brewer, 2002), for instance, or identities that involve intimacy groups, task groups, social categories, and loose associations (Lickel et al., 2000). Eliminating family, friends, and other potential sources of tangible, informational, or emotional support from the examples used by the experimenter would help to disambiguate the present results, as would eliminating (vs. emphasizing) nodes that suggest successful female role models (McIntyre et al., 2003).

Perceived Self-Efficacy

Another interesting possibility is that women high in math identification benefited from drawing detailed self-concept maps because doing so increased their perceived self-efficacy (Bandura, 1986). In the process of depicting their many social roles and identities, it might have become salient to these women that they must be very capable, because they were able to fill simultaneously so many life roles, many of them with competing demands. Thus their concerns about performing poorly on this one math test might have come to seem insignificant in comparison with their obvious multitasking skills. In a relevant study, students who were given high efficacy expectancies performed better on a difficult test when they anticipated an opportunity for self-evaluation than when they did not, whereas the opposite was true for students who were given low efficacy expectancies (Sanna & Pusecker, 1994). Indeed, Bussey and Bandura (1999) suggested that the "effect of gender on mathematical performance is mediated through perceived self-efficacy" (p. 693). Perceived self-efficacy improves performance by reducing self-impairing thought processes, stress, and depression, all of which play important roles in stereotype threat effects (Keller & Daunheimer, 2003; Schmader & Johns, 2003).

Combined Mechanisms

The present results should not be interpreted to mean that stereotype threat is always caused by identity threat or that just one of these postulated mechanisms is responsible for the ameliorative effects of multiple salient identities on performance deficits. The "experience of stereotype threat, as we have noted, is likely to vary considerably from stereotype to stereotype, setting to setting, and person to person. It is almost unimaginable that an experience as variable as this could be mediated consistently through a single psychological process" (Steele et al., 2002, p. 406). Future research on the present experiment's alleviation technique needs to explore all of the possible mechanisms by which salient multiple identities might exert an ameliorative effect on the performance of women high in math identification, but it would not be surprising to find that different mechanisms prove effective for different women, depending on some combination of individual differences in the type and importance of their affective reactions to stress, self-identity complexity, perceived self-efficacy, social support networks, and other as yet unidentified factors.

CONCLUSION

Finally, for practical application, it might not be necessary for women under stereotype threat to draw self-concept maps explicitly. It might be as effective for women to sit quietly, just before an important test such as the GRE, and mentally review all the important roles they play in life and their many competent social identities. Mental self-concept maps might prove as effective as actual paper-and-pencil maps in alleviating performance deficits under stereotype threat. Pending an empirical test of such a mental self-concept map procedure, however, the results of the present experiment suggest guarded optimism that women who first think about many aspects of their self-concepts can subsequently perform as well as men do, even in situations where they are explicitly reminded of a threatening gender stereotype.

REFERENCES

Ambady, N., Paik, S. K., Steele, J., Owen-Smith, A., & Mitchell, J. P. (2003). Deflecting negative self-relevant stereotype activation: The effects of individuation. *Journal of Experimental Social Psychology*, *40,* 401–408.

Anderson, J. R. (1985). *The architecture of cognition*. Cambridge, MA: Harvard University Press.

Aronson, J., Fried, C. B., & Good, C. (2002). Reducing the effects of stereotype threat on African American college students by shaping theories of intelligence. *Journal of Experimental Social Psychology*, *38,* 113–125.

Aronson, J., Lustina, M. J., Good, C., Keough, K., Steele, C. M., & Brown, J. (1999). When White men can't do math: Necessary and sufficient factors in stereotype threat. *Journal of Experimental Social Psychology*, *35,* 29–46.

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory.* Englewood Cliffs, NJ: Prentice Hall.

Blascovich, J., Spencer, S. J., Quinn, D., & Steele, C. (2001). African Americans and high blood pressure: The role of stereotype threat. *Psychological Science*, *12,* 225–229.

Brewer, M. B. (1991). The social self: On being the same and different at the same time. *Personality and Social Psychology Bulletin*, *17,* 475–482.

Brissette, I., Scheier, M. F., & Carver C. S. (2002). The role of optimism in social network development, coping, and psychological adjustment during a life transition. *Journal of Personality and Social Psychology*, *82,* 102–111.

Brown, R. P., & Josephs, R. A. (1999). A burden of proof: Stereotype relevance and gender differences in math performance. *Journal of Personality and Social Psychology*, *76,* 246–257.

Brown, R. P., & Pinel, E. C. (2003). Stigma on my mind: Individual differences in the experience of stereotype threat. *Journal of Experimental Social Psychology*, *39,* 626–633.

Bussey, K., & Bandura, A. (1999). Social cognitive theory of gender development and differentiation. *Psychological Review*, *106,* 676–713.

Cadinu, M., Maass, A., Frigerio, S., Impagliazzo, L., & Latinotti, S. (2003). Stereotype threat: The effect of expectancy on performance. *Journal of Experimental Social Psychology*, *33,* 267–285.

Cohen, S. (1988). Psychosocial models of the role of social support in the etiology of physical disease. *Health Psychology*, *7,* 269–297.

Conway, M., Carroll, J. M., Pushkar, D., & Arbuckle, T. (1996). Anticipated interaction, individual differences in attentional resources, and elaboration of behavior. *Social Cognition*, *14,* 338–366.

Crocker, J., Luhtanen, R. K., Cooper, M. L., & Bouvrette, A. (2003). Contingencies of self-worth in college students: Theory and measurement. *Journal of Personality and Social Psychology*, *85,* 894–908.

Croizet, J., & Claire, T. (1998). Extending the concept of stereotype threat to social class: The intellectual underperformance of students from low socioeconomic backgrounds. *Personality and Social Psychology Bulletin*, *6,* 588–594.

Davies, P. G., Spencer, S. J., Quinn, D. M., & Gerhardstein, R. (2002). Consuming images: How television commercials that elicit stereotype threat can restrain women academically and professionally. *Personality and Social Psychology Bulletin*, *28,* 1615–1628.

Ford, T. E., Ferguson, M. A., Brooks, J. L., & Hagadone, K. M. (2004). Coping sense of humor reduces effects of stereotype threat on women's math performance. *Personality and Social Psychology*, *30,* 643–653.

Gonzales, P. M., Blanton, H., & Williams, K. J. (2002). The effects of stereotype threat and double-minority status on the test performance of women. *Personality and Social Psychology Bulletin*, *28,* 659–670.

Gove, W. (1972). The relationship between sex roles, marital status, and mental illness. *Social Forces*, *51,* 34–44.

Holley, C. D., & Dansereau, D. F. (1984). *Spatial learning strategies: Techniques, applications, and related issues.* San Diego, CA: Academic Press.

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.

Inzlicht, M., & Ben-Zeev, T. (2003). Do high-achieving females underperform in private? The implications of threatening environments on intellectual processing. *Journal of Educational Processing*, *95,* 796–805.

Keller, J. (2002). Blatant stereotype threat and women's math performance: Self-handicapping as a strategic means to cope with obtrusive negative performance expectations. *Sex Roles*, *47,* 193–198.

Keller, J., & Dauenheimer, D. (2003). Stereotype threat in the classroom: Dejection mediates the disrupting effect on women's math performance. *Personality and Social Psychology Bulletin*, *29,* 371–381.

Lazarus, R. S., & Folkman S. (1984). *Stress, appraisal, and coping.* New York: Springer.

Leyens, J. Ph., Désert, M., Croizet, J.-C., & Darcis, C. (2000). Stereotype threat: Are lower status and history of stigmatization preconditions for stereotype threat? *Personality and Social Psychology Bulletin*, *26,* 1189–1199.

Lickel, B., Hamilton, D. L., Wieczorkowska, G., Lewis, A., Sherman, S. J., & Uhles, A. N. (2000). Varieties of groups and the perception of group entitativity. *Journal of Personality and Social Psychology*, *78,* 223–246.

Linville, P. W. (1985). Self-complexity and affective extremity: Don't put all of your eggs into one cognitive basket. *Social Cognition*, *3,* 94–120.

Linville, P. W. (1987). Self-complexity as a cognitive buffer against stress-related illness and depression. *Journal of Personality and Social Psychology*, *52,* 663–676.

Linville, P. W., & Carlston, D. E. (1994). Social cognition of the self. In P. G. Devine & D. L. Hamilton (Eds.), *Social cognition: Impact on social psychology* (pp. 143–193). San Diego, CA: Academic Press.

Linville, P. W., & Jones, E. E. (1980). Polarized appraisals of outgroup members. *Journal of Personality and Social Psychology*, *38,* 689–703.

Major, B., Spencer, S., Schmader, T., Wolfe, C., & Crocker, J. (1998). Coping with negative stereotypes about intellectual performance: The role of psychological disengagement. *Personality and Social Psychology Bulletin*, *24,* 34–50.

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

McGuire, W. J., & McGuire, C. V. (1988). Content and process in the experience of the self. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 21, pp. 97–144). San Diego, CA: Academic Press.

McIntyre, R. B., Paulson, R. M., & Lord, C. G. (2003). Alleviating women's mathematics stereotype threat through salience of group achievements. *Journal of Experimental Social Psychology*, *39,* 83–90.

Mickelson, K. D. (2001). Perceived stigma, social support, and depression. *Personality and Social Psychology Bulletin*, *27,* 1046–1056.

O'Brien, L. T., & Crandall, C. S. (2003). Stereotype threat and arousal: Effects on women's math performance. *Personality and Social Psychology Bulletin*, *29,* 782–789.

Pronin, E., Steele, C. M., & Ross, L. (2004). Identity bifurcation in response to stereotype threat: Women and mathematics. *Journal of Experimental Social Psychology*, *40,* 152–168.

Quinn, D. M., & Spencer, S. J. (2001). The interference of stereotype threat with women's generation of mathematical problem-solving strategies. *Journal of Social Issues*, *57,* 55–71.

Radloff, L. (1975). Sex differences in depression. *Sex Roles*, *1,* 249–265.

Rafaeli-Mor, E., & Steinberg, J. (2002). Self-complexity and well-being: A review and research synthesis. *Personality and Social Psychology Review*, *6,* 31–58.
Reevy, G. M., & Maslach, C. (2001). Use of social support: Gender and personality differences. *Sex Roles*, *44,* 437–459.
Roccas, S. (2003). Identification and status revisited: The moderating role of self-enhancement and self-transcendence values. *Personality and Social Psychology Bulletin*, *29,* 726–736.
Roccas, S., & Brewer, M. B. (2002). Social identity complexity. *Personality and Social Psychology Review*, *6,* 88–106.
Rothgerber, H. (1997). External intergroup threat as an antecedent to perceptions in in-group and out-group homogeneity. *Journal of Personality and Social Psychology*, *73,* 1206–1212.
Sanna, L. J., & Pusecker, P. A. (1994). Self-efficacy, valence of self-evaluation, and performance. *Personality and Social Psychology Bulletin*, *20,* 82–92.
Sarason, B. R., Sarason, I. G., & Gurung, R. A. R. (1997). Close personal relationships and health outcomes: A key to the role of social support. In B. R. Sarason, & S. W. Duck (Eds.), *Personal relationships: Implications for clinical and community psychology* (pp. 15–41). New York: Wiley.
Schmader, T. (2002). Gender identification moderates stereotype threat effects on women's math performance. *Journal of Experimental Social Psychology*, *38,* 194–201.
Schmader, T., & Johns, M. (2003). Converging evidence that stereotype threat reduces working memory capacity. *Journal of Personality and Social Psychology*, *85,* 440–452.
Schmader, T., Johns, M., & Barquissau, M. (2004). The costs of accepting gender differences: The role of stereotype endorsement in women's experience in the math domain. *Sex Roles*, *50,* 835–850.
Seeman, T. E. (1996). Social ties and health: The benefits of social integration. *Annals of Epidemiology*, *6,* 442–451.
Shih, M., Pittinsky, T. L., & Ambady, N. (1999). Stereotype susceptibility: Identity salience and shifts in quantitative performance. *Psychological Science*, *10,* 80–83.
Smith, J. L., & White, P. H. (2001). Development of the domain identification measure: A tool for investigating stereotype threat effects. *Educational and Psychological Measurement*, *61,* 1040–1057.
Smith, J. L., & White, P. H. (2002). An examination of implicitly activated, explicitly activated, and nullified stereotypes on mathematical performance: It's not just a woman's issue. *Sex Roles*, *47,* 179–191.
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. A., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology*, *69,* 797–811.
Steele, C. M. (1997). A threat in the air: How stereotypes shape intellectual identity and performance. *American Psychologist*, *52,* 613–629.
Steele, C. M. (1998). Stereotyping and its threat are real. *American Psychologist*, *53,* 680–681.
Steele, C. M., Spencer, S. J., & Aronson, J. (2002). Contending with group image: The psychology of stereotype and social identity threat. In M. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 34, pp. 379–440). San Diego, CA: Academic Press.
Stone, J., Lynch, C. I., Sjomeling, M., & Darley, J. M. (1999). Stereotype threat effects on Black and White athletic performance. *Journal of Personality and Social Psychology*, *77*, 1213–1227.
Stricker, L. J., & Ward, W. C. (2004). Stereotype threat, inquiring about test takers' ethnicity and gender, and standardized test performance. *Journal of Applied Social Psychology*, *34,* 665–693.
Stryker, S., & Statham, A. (1985). Symbolic interaction and role theory. In G. Lindzey & E. Aronson (Eds.), *The handbook of social psychology* (pp. 311–378). New York: Random House.
Tajfel, H. (1978). Interindividual behavior and intergroup behavior. In H. Tajfel (Ed.), *Differentiation between social groups: Studies in the social psychology of group relations* (pp. 27–60). London: Academic Press.
Taylor, S. E., Sherman, D. K., Kim, H. S., Jareho, J., Takagi, K., & Dunagan M. S. (2004). Culture and social support: Who seeks it and why? *Journal of Personality and Social Psychology*, *87,* 354–362.
Thoits, P. A. (1983). Multiple identities and psychological well-being: A reformulation and test of the social isolation hypothesis. *American Sociological Review*, *48,* 174–187.
Turner, J. C. (1987). *Rediscovering the social group: A self-categorization theory*. Oxford: Blackwell.
Walsh, M., Hickey, C., & Duffy, J. (1999). Influence of item content and stereotype situation on gender differences in mathematical problem solving. *Sex Roles*, *41,* 219–240.
Wegner, D. M., & Wenzlaff, R. M. (1996). Mental control. In E. T. Higgins & A. W. Kruglanski (Eds.), *Social psychology: Handbook of basic principles* (pp. 466–492). New York: Guilford Press.