

Research Article

Why Do Women Underperform Under Stereotype Threat?

Evidence for the Role of Negative Thinking

Mara Cadinu, Anne Maass, Alessandra Rosabianca, and Jeff Kiesner

University of Padova, Padova, Italy

ABSTRACT—*This study investigated the role of negative thinking as a potential mediator of performance deficits under stereotype threat. After being assigned to a stereotype-threat or a no-threat condition, 60 female participants were asked to complete a difficult math task. Using the thought-listing technique, women under stereotype threat reported a higher number of negative thoughts specifically related to the test and to mathematics compared with women in the no-threat condition. Moreover, women under stereotype threat also showed a sharp decrease in performance that (a) was most pronounced in the second half of the test and (b) was mediated by the increase in negative thinking.*

In the medical sciences, the first documented case of a woman scientist in Europe appears to be Trocta (or Trotula) de Ruggiero, who taught in the early 10th century at the then internationally famous Salerno Medical School in southern Italy. Seven centuries later, in 1678, Elena Cornaro Piscopia was the first woman in the world to obtain a college degree, receiving a baccalaureate in philosophy from the University of Padova, Italy. But another 190 years passed before Sofja Kovalevskaja became, in 1875, the first woman graduate in mathematics. Over the centuries, prejudice against women regarding their presumed lack of propensity for scientific disciplines in general, and mathematics in particular, has been the dominant view. Of the two most famous female mathematicians of the 19th century, their illustrious colleague Hermann Weyl (1885–1955) said: “Only two female mathematicians in history: Sofja Kovalevskaja and Emmy Noether. The first was not a mathematician, the second was not a woman” (cited in Lolli, 2000, p. 11).

Address correspondence to Mara Cadinu, Dipartimento di Psicologia dello Sviluppo e della Socializzazione, Università di Padova, via Venezia 8, 35131 Padova, Italy; e-mail: mara.cadinu@unipd.it.

Over the years the situation has clearly changed, but the stereotype that women are less talented at mathematics than men is still common. The relevant question is, then, what are the effects of these preconceptions on women themselves? In 1954, Allport suggested that members of stigmatized minority groups might inadvertently confirm what majority-group members expect of them. Although little attention was paid to this issue in psychological research for several decades, a renewed interest in the topic can be found in work by Crocker and Major (1989), Steele and his collaborators (Spencer, Steele, & Quinn, 1999; Steele & Aronson, 1995), Swim and Stangor (1998), and other investigators.

According to the recent model of stereotype threat, proposed by Steele and his collaborators, the activation of a negative stereotype regarding a minority group will interfere with the performance of group members in stereotype-relevant domains. For example, if the stereotype that women are unskilled in mathematics becomes salient in a testing situation, this stereotype activation will divert women’s attention onto task-irrelevant worries and induce anxiety caused by the fear of confirming the negative stereotype. According to this model, increased anxiety will then reduce cognitive resources, thus leading to performance deficits. In the long run, this process may also lead minority members to “disidentify” with the stereotype-relevant domain. Studies consistent with the stereotype-threat model have documented performance deficits under stereotype threat for several minority groups, including ethnic minorities (e.g., Steele & Aronson, 1995), women (e.g., Spencer et al., 1999), girls (e.g., Ambady, Shih, Kim, & Pittinsky, 2000), students from low socioeconomic backgrounds (Croizet & Claire, 1998), and old people (Levy, 1996).

Although there is little doubt about the pervasiveness of the phenomenon, it is striking that relatively little is known about the processes underlying performance deficits under stereotype threat. The first explanation, originally proposed by Steele and Aronson (1995), suggests that minority-group members are afraid

to confirm the stereotype that associates their own group with poor performance in a given domain. In addition to “normal,” or baseline, anxiety associated with taking difficult tests, minority-group members may therefore experience further tension due to their preoccupation with confirming the negative stereotype. Because high levels of anxiety and arousal have been found to be detrimental to performance on difficult tasks, it is plausible that stereotype-induced increases in anxiety interfere with performance (see Hill & Wigfield, 1984; see Dembo & Eaton, 1997, for an overview). However, contrary to this prediction, work on stereotype threat has found null or weak effects of stereotype threat on anxiety.

ANXIETY

Steele and Aronson (1995, Experiment 2) measured participants' levels of anxiety immediately after a test and found no effects of stereotype threat on anxiety. Aronson et al. (1999) also found no effects of stereotype threat on a measure of anxiety taken at the end of a test. Stone, Lynch, Sjomeling, and Darley (1999, Experiment 1) assessed self-reported situational anxiety in their Experiment 1 and found a nonsignificant tendency in the predicted direction, but anxiety was unrelated to performance. In a subsequent study (Experiment 2), they measured state anxiety before and after the test and found a general increase in anxiety, but this increase was no stronger in the stereotype-threat than in the control condition. More encouraging is a study by Spencer et al. (1999, Experiment 3), who found that stereotype threat had a marginal effect on anxiety, which in turn had a negative effect on performance. However, further analyses failed to support the mediational role of anxiety on performance. In a more recent study, Osborne (2001) found that race differences in test performance were mediated by anxiety. However, no strong conclusions can be drawn from this study because the mediator (anxiety) was measured after the outcome variable (performance). Overall, despite the numerous attempts to test the role of anxiety in performance under stereotype threat, the empirical evidence suggesting that anxiety plays a causal role in stereotype-threat performance decrements is quite limited at this point.

The limited evidence may, in part, reflect a measurement problem. Most anxiety measures used in stereotype-threat research are scales such as the State-Trait Anxiety Inventory (Spielberger, Gorsuch, & Lushene, 1970), which include items like “I am worried,” “I feel nervous,” and “I am jittery.” However, these scales may not be suited to tap those facets of anxiety that are most closely linked to domain-specific performance. More specifically, the literature on anxiety identifies two distinct components of anxiety: (a) emotionality, referring to people's awareness of bodily arousal and tension (Deffenbacher, 1977; Kaplan, McCordick, & Twitchell, 1979; Liebert & Morris, 1967; Morris, Davis, & Hutchings, 1981), and (b) worry, referring to the cognitive side of anxiety (preoccupations, concerns). Research has demonstrated that the performance-debilitating effects of

anxiety are specifically related to the worry component, whereas the emotionality component is unrelated to performance decrements (Deffenbacher, 1980; Tyron, 1980). Similarly, research by Deffenbacher and Deitz (1978) has suggested that cognitive interference is the key predictor of performance deficits (see Sarason, 1984, for similar findings).

THOUGHT INTRUSION

Unfortunately, studies focusing on intrusive thoughts, rather than anxiety, are rare in the stereotype-threat literature, and most evidence is indirect at best. Indirect evidence in line with the hypothesis that performance deficits under stereotype threat are due to intrusive thoughts comes from a study by Schmader and Johns (2003), who found a reduction of working memory capacity, which in turn led to lowered performance on a math test, among women exposed to stereotype threat. This result is in line with the idea that stigmatized individuals under stereotype threat have interfering thoughts that reduce their working memory capacity. However, because no thought measure was taken in this study, it remains unclear what kind of disruptive thoughts were competing for cognitive resources.

Studies that have employed thought measures have produced only weak (and inconsistent) evidence for thought intrusion under stereotype threat. In two studies in which a stronger decrease in performance was found in the stereotype-threat condition than in the control condition (Steele & Aronson, 1995, Experiments 1 and 2), the authors included among their postexperimental measures a cognitive interference scale, assessing disruptive thoughts and feelings during the test. However, no effects of condition were found on the cognitive-interference measure. In Steele and Aronson's third experiment, they found evidence for both stereotype activation and self-doubt on a word-completion task, but because no performance measure was taken in this study, the mediating role of these variables remains unknown. More recently, Stone (2002), in two independent studies, employed a word-completion task assessing the activation of thoughts related to race, self-doubt, and athleticism before a series of practice trials, but found an increase in stereotype-threat-related activation on only 2 out of 17 items. More important, none of these studies reporting intrusive thoughts under stereotype threat assessed actual performance, thus making it impossible to test for mediation.

Thus, although the intrusive-thoughts hypothesis figures among the most promising explanations, there is no empirical proof so far that intrusive thoughts do, indeed, mediate performance deficits under stereotype threat. Moreover, all prior studies have involved reactive-thought measures (either scales or word completion) whose content is predetermined by the experimenter, and that may be quite different from the thoughts that stigmatized individuals generate spontaneously under stereotype threat.

THE PRESENT STUDY

For these reasons, in the present work we focused on the interfering thoughts that participants spontaneously generate during a task in a stereotype-threat condition versus a no-threat condition. Thus, we focused not on general test anxiety, but rather on specific intrusive thoughts that have been found to play a central role in anxiety-related cognitive deficits (Sarason, 1984). Specifically, we predicted that intrusive thoughts, such as a preoccupation regarding one's performance or a sense of inadequacy regarding the tested domain, should occur more frequently under stereotype threat than in a no-threat (control) condition. To test this hypothesis, we used the thought-listing technique to measure the spontaneous occurrence of such thoughts. Further, we expected that the increase in negative thoughts during the task would be associated with a decrease in task performance, as suggested by the intrusive-thoughts hypothesis. In other words, we expected stereotype threat to lead to negative thinking, which in turn would lead to a decrease in performance. Because negative thoughts are believed to mediate the effects of stereotype threat, we expected to find an immediate effect of stereotype threat on negative thoughts. However, it is possible that the effects of stereotype threat on performance do not occur immediately after the threat, but are delayed. If the cognitive-interference hypothesis is correct, then the debilitating effects of negative thoughts could be most pronounced on the later trials of the task, because of the cumulative effects of thought intrusions. Therefore, we also tested for a temporal ordering of the effects. Thus, the main goal of the present study was to test the mediational role of intrusive thoughts in stereotype-threat performance deficits.

To test these hypotheses, we asked a group of women to complete a difficult mathematics test under stereotype threat or in a no-threat (control) condition. During the task, the thought-listing technique (Krohne, Pieper, Knoll, & Breimer, 2002; Petty & Cacioppo, 1977, 1979) was used to measure participants' spontaneous thoughts. We predicted that (a) individuals under stereotype threat would report a higher number of negative thoughts regarding the math domain and their perceived lack of competence on the math test compared with individuals in the control condition; (b) individuals under stereotype threat would show a decrease in performance compared with individuals in the control condition; (c) negative thoughts would mediate the negative effects of stereotype threat on performance; and (d) the occurrence of negative thoughts would be observed immediately following the stereotype threat, whereas the performance deficit might occur following a lag of several trials (e.g., as negative thoughts accumulated).

METHOD

Participants

Sixty female psychology students at the University of Padova, Italy, volunteered for this study.

Experimental Manipulation

Participants in the stereotype-threat condition were told that "recent research has shown that there are clear differences in the scores obtained by men and women in logical-mathematical tasks." Control participants were told that "there are no differences between men and women in logical-mathematical tasks."

Mathematics Test and Thought Listing

The material for the experiment consisted of a difficult mathematics test composed of seven problems similar to those on the Graduate Record Examination (average number of correct responses across conditions = 4.4, $SD = 1.77$). Before each exercise there was a blank page with the following instruction: "Now, please write anything that comes to your mind." Before beginning the experiment, participants had been told that before each exercise they would "find an empty page on which they could write anything they wanted, for example, 'Tomorrow is my mother's birthday.'"

The thought-listing sentences were coded by two independent judges (interrater agreement = 91%) using the following categories: (a) *negative math-related thoughts*, including thoughts like "These exercises are too difficult for me" and "I am not good at math"; (b) *hate for math*, including thoughts like "I have always hated math" and "I hate math"; (c) *generic distress*, including thoughts like "I am tired" and "It is hot here"; (d) *self-confidence*, including thoughts like "I am really good" and "I love numbers"; (e) *neutral reference to the test*, including thoughts like "Probably none of these alternatives are correct" and "How do we know that d and g are the minimal values?"; (f) *not knowing what to write*, including thoughts like "I don't know what to write" and "Nothing"; (g) *guess*, including thoughts like "For this one I had no clue" and "I guessed"; and (h) *other*, including thoughts like "pizza" and "Mario, I love you."

RESULTS

Performance

As in previous studies, a significantly lower number of correct responses was found in the stereotype-threat condition ($M = 3.93$) than in the control condition ($M = 4.87$), $t(58) = 2.10$, $p < .05$, Cohen's $d = 0.55$.¹ To test whether performance deficits were most pronounced in the second half of the test, we conducted an analysis of variance (ANOVA) on the number of correct responses, with experimental condition (threat vs. no-threat) as a between-participants factor and time (first three problems vs. last

¹This difference cannot be accounted for by disengagement because practically all participants attempted to resolve all problems. Cases in which participants did not attempt to resolve a problem were extremely rare (3% in the control group and 5% in the experimental group).

three problems) as a within-participants factor.² As predicted, a highly significant interaction was found, $F(1, 58) = 8.69, p < .01$. During the first part of the test, there was no difference between participants under threat ($M = 2.07$) and participants in the no-threat condition ($M = 2.10$), whereas in the second part of the test, stereotype-threat participants showed a significant performance deficit ($M = 1.67$) compared with no-threat participants ($M = 2.43$), $t(58) = 3.1, p < .01$, Cohen's $d = 0.81$. It should be noted that, whereas the control group showed a marginally significant increase in performance from the first three problems to the last three problems, $t(29) = 1.84, p < .08$, Cohen's $d = 0.68$, the experimental group showed a significant decrease in performance, $t(29) = 2.35, p < .05$, Cohen's $d = 0.87$. Note that these findings cannot be accounted for by greater difficulty of the second half of the test. Looking at the correct responses in the control condition, it becomes clear that the last three items (87%, 77%, and 80% correct) were, if anything, slightly easier than the first three items (53%, 87%, and 70% correct). Moreover, examining the performance differential, by subtracting for each trial the percentage of correct responses in the stereotype-threat condition from the percentage of correct responses in the control condition (see Fig. 1), it becomes clear that the performance deficit of the experimental (compared with control) participants occurred only after the first trial, and was marked by a steady decline across the remainder of the trials.

Thought Listing

We predicted that participants under stereotype threat would report more negative thoughts related to the test or the math domain than participants in the control condition. Because the thought-listing distribution was positively skewed, we performed a log-transformation on the thought-listing scores.³ Results were consistent with the prediction. Participants in the stereotype-threat condition showed a significantly higher number of negative math-related thoughts ($M = 1.47$) than participants in the control condition ($M = 0.60$), $t(58) = 2.55, p < .05$, Cohen's $d = 0.67$. Also consistent with predictions was the finding that the level of negative thinking was similar for the first and second halves of the test. In a 2×2 ANOVA including experimental condition as a between-participants factor and time (first three thoughts vs. last three thoughts) as a within-participants factor, no interaction between these variables was found, $F(1, 58) = 0.034$, n.s. The hate-for-math thought category had a very low base rate ($M = 0.07$) and was unaffected by the stereotype-threat manipulation, a result that probably reflects extreme and stable attitudes toward math that did not vary as a function of the

²We did not include Problem 4 in these analyses in order to have a raw-score common metric (sum of three exercises) for the repeated measures score. However, additional analyses showed that the results were the same if an average score was used as the dependent variable and Problem 4 was included in the analyses, either in the first half or in the second half.

³The skewness (γ) for the nontransformed thought-listing scores was 1.32, whereas the skewness for the transformed thought-listing scores was 0.55.

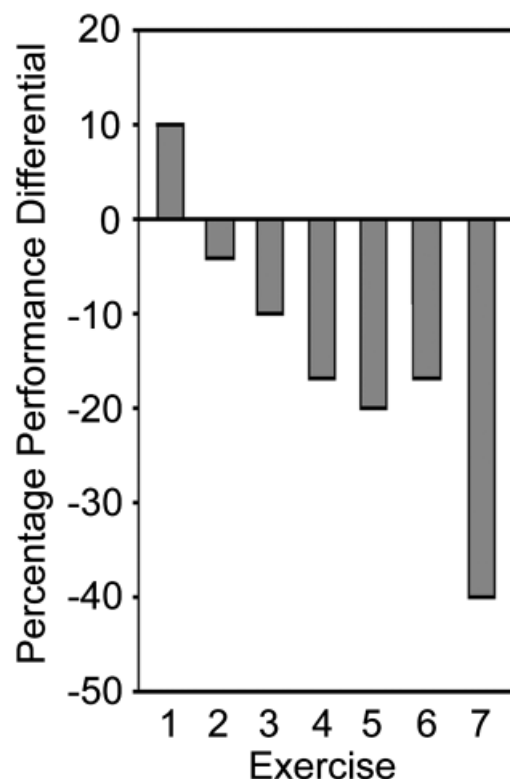


Fig. 1. Performance differentials (percentage of correct responses in the stereotype-threat condition minus percentage of correct responses in the control condition). Positive values indicate better performance under stereotype threat; negative values indicate better performance in the control condition.

manipulation. No effects were found on any of the other thought-listing categories. In particular, for generic distress, no differences were found between the stereotype-threat condition ($M = 0.23$) and the control condition ($M = 0.43$), $t(58) = 1.23$, n.s. Therefore, in the remainder of this article, we use the term *thought-listing* specifically to refer to reports of negative math-related thoughts.

The Mediating Role of Thought Listing on Performance

We hypothesized that negative math-related thoughts would mediate the effect of the stereotype-threat manipulation on performance. To test for direct and indirect effects of stereotype threat on performance and thought listing, we used structural equation modeling (LISREL 8.54; Jöreskog & Sörbom, 1996). The specific model used in this analysis is presented in Figure 2. This model allows the stereotype-threat manipulation to have direct effects on thought listing and performance at both Time 1 (Exercises 1–3) and Time 2 (Exercises 4–7) and indirect effects on thought listing and performance at Time 2. Additionally, given our prediction that indirect effects on performance would be mediated by thought listing, we allowed a path from Time 2 thought listing to Time 2 performance.

Results indicated that the overall model fit the data well, $\chi^2(1) = 0.15, p < .70$, root mean square error of approximation (RMSEA) = .00. An examination of the regression coefficients

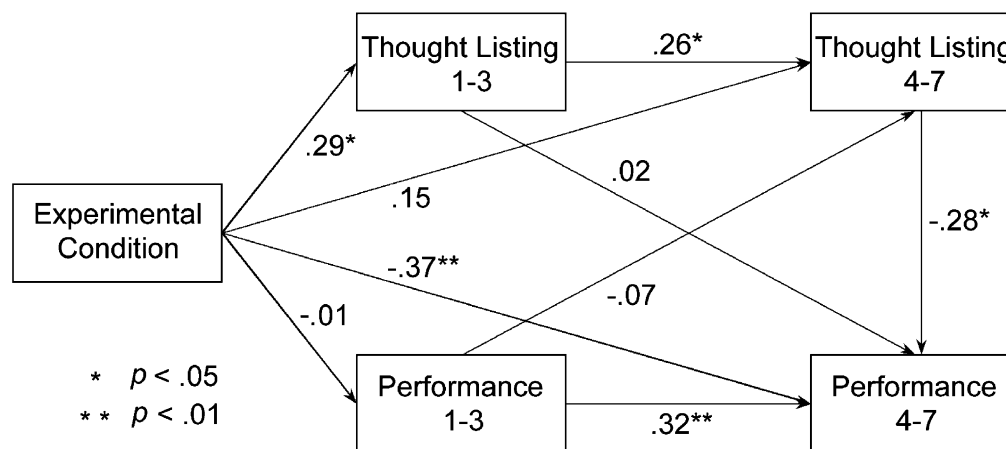


Fig. 2. Structural equation model predicting effects of stereotype threat on thought listing and performance for Problems 1 through 3 and Problems 4 through 7.

indicated that, as seen in the ANOVA, stereotype threat had an immediate effect on thought listing, but no immediate effect on performance. Moreover, both thought listing and performance showed significant levels of stability from Time 1 to Time 2. Finally, and most important, this model provides support for the hypothesis that stereotype threat has an indirect effect on performance that is mediated through thought listing. Specifically, stereotype threat had an effect on Time 1 thought listing, which significantly predicted Time 2 thought listing, which in turn significantly predicted Time 2 performance. Note that the standardized path coefficient for the main effect of stereotype threat on Time 2 performance without controlling for thought listing was $-.43$, as compared with $-.37$ in the analysis presented in Figure 2. Thus, this model supports the hypothesis that negative math-related thought intrusions mediate the effects of stereotype threat on performance.

To further test our hypothesis of mediation, we tested the same model eliminating the path from Time 2 thought listing to Time 2 performance. This reduced model no longer fit the data well, $\chi^2(2) = 6.32, p < .05$, RMSEA = $.19$, and the fit was significantly reduced from that of the full model, $\Delta\chi^2(1) = 6.17, p < .05$. These results indicate that allowing this path, and thus allowing for the mediational effect of thought listing, is important for having a good-fitting model.

DISCUSSION

The goal of the present study was to investigate one mechanism underlying stereotype-related performance deficits that has received surprisingly little attention in previous research. In the present study, women were presented with a difficult mathematics test either under stereotype threat or in a no-threat condition. As predicted, women in the stereotype-threat condition underperformed compared with women in the control condition. Most important, the decrease in performance was mediated by the increase in negative domain-specific thinking. More specific-

ly, compared with women in the control condition, women under stereotype threat reported a higher number of negative thoughts regarding their inability to perform the test and their lack of competence in math. The mediational analysis supported the hypothesis that stereotype threat led to negative thoughts, which in turn led to a temporally lagged decrease in performance.

The present findings are the first to demonstrate a link between negative thinking and performance deficits. Thus, they are the first to provide direct evidence in support of Steele and his collaborators' hypothesis (Steele & Aronson, 1995) that performance deficits under stereotype threat are caused by intrusive thoughts that occur during task performance. These results are also in line with recent research by Schmader and Johns (2003), who showed reduction in working memory capacity under stereotype threat, but further specify that it is individuals' domain-specific disruptive thoughts that have a detrimental effect on performance. More specifically, by using the thought-listing technique, the present study demonstrated that participants under stereotype threat spontaneously engage in negative task-related thinking and that this negative thinking inhibits subsequent performance. Our data also suggest that performance deficits under stereotype threat are caused by domain-specific thought intrusion, and not by general discomfort or anxiety.

A second important finding of this study is that the effects for thought listing occurred immediately after the manipulation, but the effects for performance were temporally delayed. This suggests that the manipulation had an immediate effect on negative thoughts, which then had cumulative negative effects on performance. This result has two important implications. First, this temporal pattern provides strong evidence for the hypothesis that negative thoughts are the cause of poor performance under stereotype threat, rather than being a consequence of the poor performance. However, future research may extend this reasoning by investigating the possibility that the two variables interact in a dynamic and mutually reinforcing way. Thus, intrusive thoughts at the beginning of the test may interfere with subse-

quent performance, which may in turn increase the likelihood of negative thinking, ultimately leading to an accelerating, self-maintaining cycle. Second, the temporal pattern observed has implications for future research, suggesting that it may be crucial to investigate the temporal pattern of performance deficits under threat. If the negative effects of stereotype threat become more evident as a test progresses, previous research may have underestimated the strength of stereotype threat.

Stereotype threat also had a direct effect on Time 2 performance, even after controlling for the indirect effect through thought listing. There are three plausible explanations for this result. First, there simply may be a delayed effect of stereotype threat on performance. However, this explanation is unlikely to be correct because a delayed effect without mediation would be difficult to explain. A second and more plausible explanation is that thought listing fully mediates the effect of stereotype threat on performance, although in the present data thought listing did not provide full mediation. We propose that thought listing could provide such full mediation, but that the effects of thought listing were attenuated because of its low base rate. One important limitation of the thought-listing measure is that it relies on spontaneous disclosure of negative test-related intrusive thoughts. For the purpose of this study, this was the only acceptable approach to measurement because prompting or asking specifically about negative test-related thoughts would have created an additional intervention. However, this measure of thought listing likely underestimates the actual level of negative test-related thinking. If so, then the effect of thought listing would be attenuated, as would be the mediational role of thought listing. Thus, it is possible that negative intrusive thoughts do fully mediate the effects of stereotype threat on performance, but that the measure of negative thoughts we used does not capture the full variance of this variable, resulting in attenuated mediation effects. The third possible explanation is that thought listing may indeed be a partial mediator of performance, and there may be additional mediators (e.g., expectancies) not assessed in the present study that contributed independently to the observed performance deficits.

To conclude, the results from the present study represent a significant step toward understanding the ubiquitous performance deficits of individuals under stereotype threat by demonstrating the mediating role of negative thinking as a crucial factor contributing to the decrease in performance. Although thought intrusion may not be the only process involved, our data suggest that it is an important and significant mediator of stereotype-threat performance deficits. This is an important step considering that modern society is faced with the complex task of reducing the barriers toward minority-group members' success.

Acknowledgments—The authors thank Luciano Arcuri, Luigi Castelli, and Jeroen Vaes for their helpful suggestions and Laura Scandolo for her help in collecting data.

REFERENCES

- Allport, G.W. (1954). *The nature of prejudice*. Reading, MA: Addison-Wesley.
- Ambady, N., Shih, M., Kim, A., & Pittinsky, T.L. (2001). Stereotype susceptibility in children: Effects of identity activation on quantitative performance. *Psychological Science, 12*, 385–390.
- 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.
- Crocker, J., & Major, B. (1989). Social stigma and self-esteem: The self-protective properties of stigma. *Psychological Review, 96*, 608–630.
- Croizet, J.C., & 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, 24*, 588–594.
- Deffenbacher, J.L. (1977). Relationship of worry and emotionality to performance on the Miller Analogies Test. *Journal of Educational Psychology, 69*, 191–195.
- Deffenbacher, J.L. (1980). Worry and emotionality in test anxiety. In G. Sarason (Ed.), *Test anxiety: Theory, research, and applications* (pp. 11–128). Hillsdale, NJ: Erlbaum.
- Deffenbacher, J.L., & Deitz, S.R. (1978). Effects of test anxiety on performance, worry, and emotionality in naturally occurring exams. *Psychology in the Schools, 15*, 446–450.
- Dembo, M.H., & Eaton, M.J. (1997). School learning and motivation. In G.D. Pyle (Ed.), *Handbook of academic learning: Construction of knowledge* (pp. 65–103). San Diego, CA: Academic Press.
- Hill, K.T., & Wigfield, A. (1984). Test anxiety: A major educational problem and what can we do about it. *Elementary School Journal, 85*, 105–126.
- Jöreskog, K.G., & Sörbom, D. (1996). *LISREL 8: User's reference guide*. Chicago: Scientific Software.
- Kaplan, R.M., McCordick, S.M., & Twitchell, M. (1979). Is it the cognitive or the behavioral component which makes cognitive-behavior modification effective in test anxiety? *Journal of Counseling Psychology, 26*, 371–377.
- Krohne, H.W., Pieper, M., Knoll, N., & Breimer, N. (2002). The cognitive regulation of emotions: The role of success vs. failure experience and coping dispositions. *Cognition and Emotion, 16*, 217–243.
- Levy, B. (1996). Improving memory in old age through implicit self-stereotyping. *Journal of Personality and Social Psychology, 71*, 1092–1107.
- Liebert, R.M., & Morris, L.W. (1967). Cognitive and emotional components in test anxiety: A distinction and some initial data. *Psychological Reports, 20*, 975–978.
- Lolli, G. (2000). *La crisalide e la farfalla: Donne e matematica*. Torino, Italy: Bollati Boringhieri.
- Morris, L.W., Davis, M.A., & Hutchings, C.H. (1981). Cognitive and emotional components of anxiety: Literature review and a revised Worry-Emotionality Scale. *Journal of Educational Psychology, 73*, 541–555.
- Osborne, J.W. (2001). Testing stereotype threat: Does anxiety explain race and sex differences in achievement? *Contemporary Educational Psychology, 26*, 291–310.
- Petty, R.E., & Cacioppo, J.T. (1977). Forewarning, cognitive responding, and resistance to persuasion. *Journal of Personality and Social Psychology, 35*, 645–655.

- Petty, R.E., & Cacioppo, J.T. (1979). Effects of forewarning of persuasive intent and involvement on cognitive responses and persuasion. *Journal of Personality and Social Psychology*, *5*, 173–176.
- Sarason, I.G. (1984). Stress, anxiety, and cognitive interference: Reactions to tests. *Journal of Personality and Social Psychology*, *46*, 929–938.
- Schmader, T., & Johns, M. (2003). Converging evidence that stereotype threat reduces working memory capacity. *Journal of Personality and Social Psychology*, *85*, 440–452.
- 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.
- Spielberger, C.D., Gorsuch, R.L., & Lushene, R.E. (1970). *Manual for the State-Trait Anxiety Inventory*. Palo Alto, CA: Consulting Psychologists Press.
- 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.
- Stone, J. (2002). Battling doubt by avoiding practice: The effects of stereotype threat on self-handicapping in White athletes. *Personality and Social Psychology Bulletin*, *28*, 1667–1678.
- 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.
- Swim, J.K., & Stangor, C. (1998). *Prejudice: The target's perspective*. San Diego, CA: Academic Press.
- Tyron, G.S. (1980). The measurement and treatment of test anxiety. *Review of Educational Research*, *50*, 343–372.

(RECEIVED 2/25/04; REVISION ACCEPTED 1/20/05)