
A Metacognitive Perspective on the Cognitive Deficits Experienced in Intellectually Threatening Environments

Toni Schmader
Chad E. Forbes
Shen Zhang
University of Arizona

Wendy Berry Mendes
Harvard University

Three studies tested the hypothesis that negative metacognitive interpretations of anxious arousal under stereotype threat create cognitive deficits in intellectually threatening environments. Study 1 showed that among minority and White undergraduates, anxiety about an intelligence test predicted lower working memory when participants were primed with doubt as compared to confidence. Study 2 replicated this pattern with women and showed it to be unique to intellectually threatening environments. Study 3 used emotional reappraisal as an individual difference measure of the tendency to metacognitively reinterpret negative emotions and found that when sympathetic activation was high (indexed by salivary alpha-amylase), women who tended to reappraise negative feelings performed better in math and felt less self-doubt than those low in reappraisal. Overall, findings highlight how metacognitive interpretations of affect can undermine cognitive efficiency under stereotype threat and offer implications for the situational and individual difference variables that buffer people from these effects.

Keywords: *stereotype threat; test performance; working memory; anxiety; sympathetic activation; emotional reappraisal*

Imagine two students taking the Graduate Record Exam (GRE). Both experience nervous energy as they begin the test. However, one student feels confident about his performance and interprets his arousal as a sign that the test is a challenge. The other is plagued by self-doubt and assumes that her arousal signals impending failure. Although they might have equivalent ability, the latter

student's tendency to metacognitively interpret her nerves in light of self-doubt could interfere with cognitive processing and impair her performance. The present article reports three studies that test how such metacognitive processes function among students likely to experience self-doubt in intellectual testing situations due to stereotype threat.

Research has demonstrated that subtle reminders of being negatively stereotyped can impair performance on complex cognitive tasks (Steele & Aronson, 1995; Steele, Spencer, & Aronson, 2002). Although the processes underlying these impairments remained enigmatic for nearly a decade, theory and research has now identified a host of factors that work individually or in concert to reduce the performance of stereotyped targets. In a recent process model of the mechanisms underlying stereotype threat, Schmader, Johns, and Forbes (2008) argue that stereotype threat leads individuals to monitor the situation for signs that they might be confirming the

Authors' Note: This research was supported by National Institute of Mental Health Grant 1R01MH071749 awarded to the first author. Many thanks to Meghan Anaya, Paul Apicella, Ashley Bunch, Alyssa Caraway, Lisa Farrell, Dan Fisher, Jared Garner, Chelsea McConnell, Rachel Nathanson, Nicole Pugliese, and David Ryan who assisted with data collection, and to Michael Johns for comments on an earlier draft. Correspondence should be addressed to Toni Schmader, Department of Psychology, University of Arizona, Tucson, AZ 85721; e-mail: schmader@u.arizona.edu.

Editor's Note: Dr. William von Hippel served as guest action editor for this article

PSPB, Vol. 35 No. 5, May 2009 584-596

DOI: 10.1177/0146167208330450

© 2009 by the Society for Personality and Social Psychology, Inc.

stereotype in order to avoid doing so. However, this avoidance motivation biases the stereotyped target toward interpreting ambiguous cues as potential evidence of failure. As a result, arousal that might be normal in a high-stakes performance situation is interpreted in a negative way, triggering ruminative processes that can absorb the executive resources needed for successful performance on many complex cognitive tasks.

Although these processes have not been directly tested, previous research highlights the role of affective responses in stereotype threat. For example, those who experience stereotype threat in intellectual or social contexts sometimes show evidence of increased physiological arousal (Ben-Zeev, Fein, & Inzlicht, 2005; Murphy, Steele, & Gross, 2007) or physiological threat (Blascovich, Spencer, Quinn, & Steele, 2001; Mendes, Blascovich, Lickel, & Hunter, 2002). This increased arousal resulting from stereotype threat has also been captured by implicit and, at least in some studies, explicit measures of anxiety (Bosson, Haymovitz, & Pinel, 2004; Johns, Inzlicht, & Schmader, 2008; Marx & Stapel, 2006; Spencer, Steele, & Quinn, 1999).

In addition to these affective processes, studies have highlighted negative thoughts as a mediator of threat-induced performance effects. This includes research showing that stereotype threat undermines expectancies for one's performance (Stangor, Carr, & Kiang, 1998) and activates negative thoughts and self-doubt (Beilock, Rydell, & McConnell, 2007; Steele & Aronson, 1995) that in one study mediated stereotype threat effects on performance (Cadinu, Maass, Rosabianca, & Kiesner, 2005). Although Cadinu et al. (2005) acknowledged that negative thoughts might only be partially mediating stereotype threat effects, this research showed that negative thinking plays some role in performance impairments.

Finally, research has pointed to cognitive interference as a central mechanism underlying stereotype-threat-induced impairments to performance (Schmader et al., 2008). For example, Schmader and Johns (2003) found that both women and minorities exhibited lower working memory when they anticipated that their math ability or intellectual skill, compared with a more neutral task frame, would be assessed in the study. This reduction in working memory mediated threat effects on later test performance. Subsequent studies using other types of methods have also concluded that situations of stereotype threat tax or deplete cognitive resources (Beilock et al., 2007; Croizet, Despres, Gauzins, Huguot, & Leyens, 2004; Inzlicht, McKay, & Aronson, 2006).

The preceding research points to arousal or anxiety, negative cognitions, and reduced cognitive resources as all playing some role in undermining performance under stereotype threat. However, these variables have been studied in isolation and the findings for these mediational relationships have varied. Anxiety is not always

elevated by stereotype threat and does not always predict lower performance (Wheeler & Petty, 2001). Similarly, negative thinking has not always been directly linked to poor performance (Spencer et al., 1999; Steele & Aronson, 1995). These mixed results raise the possibility that anxious arousal and negative thoughts do not directly impair performance through independent paths but rather interact to reduce the cognitive resources needed for intellectual tasks. We propose a metacognitive perspective on stereotype threat where primed cognitions moderate the meaning of affective experience in an intellectually threatening environment, altering the relationship that anxious arousal has to performance.

Researchers have recently resuscitated an interest in the role of metacognitive thinking in social psychological phenomenon (Jost, Kruglanski, & Nelson, 1998). As social information processors, we often try to understand what our psychological and physiological states might mean given our current situation. For example, emotion research shows that individuals use their affective states to judge whether something in the situation demands further attention (in the case of negative affect) or is safe for relaxation (in the case of positive affect; Schwarz & Clore, 1983). In the attitudes literature, flexing one's arms away from the body and nodding one's head are interpreted as subtle cues to avoidance and confidence, respectively, which can then reinforce one's interpretation of a self-relevant persuasive message (e.g., Briñol & Petty, 2003). In this same metacognitive vein, appraisal theories of emotion assume that the discrete emotional states we experience are a function both of physiological changes and cognitive interpretations of what those changes might indicate given the surrounding context (Lazarus, 1991; Schachter & Singer, 1962). Thus, generalized arousal can take on different psychological meaning depending on one's interpretation of what that arousal might be signaling.

Similarly, initially activated thoughts and discrete emotions can have pernicious effects on later processing if they are evaluated as being harmful. For example, believing one's intrusive thoughts and feelings are inappropriate (i.e., worrying about worrying) or uncontrollable can contribute to further negative thinking (Watkins, 2004; Wells, 1995). As suggested in the opening paragraph, individuals of equivalent ability and initial anxiety might still perform differently on a task if some interpret their anxiety in a negative way but others reappraise it more positively. Such processes should be most pronounced during self-relevant tasks that activate a motivation toward self-evaluation in the face of potential self-threat (Tesser, 1988).

Although anyone could interpret heightened arousal during a self-relevant task as a negative cue, situations that prime stereotype threat are especially prone to

engender these effects (Johns & Schmader, in press). Whereas arousal or even anxiety in a testing context could be viewed as being fairly normal, those who are negatively stereotyped and thus have a heightened accessibility of negative thoughts might be more likely to interpret their affective state as a sign of poor performance. This biased interpretation is then likely to spawn further negative ruminations that consume working memory resources that would be critical for optimal performance (Beilock et al., 2007).

If stereotype threat typically evokes negative thoughts that bias how stereotyped targets interpret their affective state, thereby depleting cognitive resources, then these effects should be circumvented by situations that prime positive thinking or for individuals who tend to reappraise negative emotions. In other words, stigmatized individuals who experience thoughts of confidence or who reappraise their arousal in a more positive way might react similarly to positively stereotyped individuals. Research on “stereotype lift” shows performance facilitation when individuals are subtly primed with a positively stereotyped social identity (Shih, Ambady, Richeson, Fujita, & Gray, 2002). Researchers have speculated that this performance boost might stem from confidence that is primed by one’s positively stereotyped status (Cheryan & Bodenhausen, 2000). As a result, the beneficiaries of positive stereotypes might typically show a pattern of performance that is unrelated to their initial affective experience. However, priming them with negative thoughts during a self-relevant task should reverse this pattern so that arousal becomes predictive of cognitive interference, even though they would not normally experience stereotype threat.

Although this metacognitive hypothesis has not been experimentally tested, there have been experiments testing interventions designed to change the way people interpret negative thoughts and feelings stemming from stereotype threat. For example, Walton and Cohen (2007) found that providing minority college students with evidence that anxiety and frustration are a normal part of coming to college led to marked improvements in grades over the ensuing months. In a laboratory context, manipulations designed to reframe anxiety as being externally caused by negative stereotypes (Johns, Schmader, & Martens, 2005) or as unrelated to performance (Johns et al., 2008) have also been effective at reducing performance impairments due to stereotype threat.

Even though these studies suggest that some kind of secondary interpretation of anxiety might be occurring during situations of stereotype threat, no prior research has directly tested the hypothesis that negative thoughts would alter one’s interpretation of an affective state to predict cognitive processing. In addition, studies that have manipulated how anxiety or arousal might be

appraised have used explicit manipulations that could affect other psychological processes. For example, Walton and Cohen (2007) suggest that their intervention promotes a sense of belonging that might both increase academic investment and reduce threat. Furthermore, the studies that have examined negative thinking as a mediator of stereotype threat effects have conceptualized this variable as having an independent relationship to lower performance (rather than being a moderator of arousal) and only measured negative thinking to test its correlation with overall performance.

In contrast to this existing research, the aim of our first two studies was to directly but unobtrusively manipulate cognitions within a performance context to test our general hypothesis that arousal experienced in an intellectually threatening environment will predict cognitive deficits when individuals are primed with doubt, but confidence primes will attenuate this relationship. Because prior stereotype threat research has most frequently measured people’s affective experiences in terms of self-reported anxiety, Studies 1 and 2 operationalized arousal as state anxiety, whereas Study 3 used a more direct measure of sympathetic activation. Building from the idea that confidence can change the relationship between anxiety and cognitive performance, Study 3 took an individual difference approach to test the hypothesis that at high levels of sympathetic activation, women with a tendency to reappraise negative emotions perform better than those who are less likely to metacognitively reinterpret their affective experiences. When sympathetic activation is low, reappraisal should have no bearing on performance.

STUDY 1

Study 1 examined the working memory of minority and White college students anticipating a diagnostic intelligence test. Previous research has established that reminders of race or intelligence testing prime thoughts of self-doubt for ethnic minority college students but not for their White peers (Steele & Aronson, 1995). If a key group difference in the experience of intellectual testing is whether those situations evoke confidence (for the positively stereotyped group) or doubt (for the negatively stereotyped group), then directly manipulating these thoughts in both White and minority students should lead the two groups to respond similarly. Specifically, we expected primed thoughts of confidence or doubt to interact with participants’ initial experience of anxiety in the situation to predict working memory. When primed with doubt, anxiety should predict lower working memory, but when primed with confidence, this relationship should be mitigated.

Method

Participants

Participants were 78 undergraduates who self-identified as White American, Hispanic American, African American, or American Indian. All participants received either course credit or \$10 for their time. Participants were considered eligible for the study if they reported on a pretest having knowledge of the stereotype that Whites are perceived to be more intelligent than stigmatized minorities in society.¹ One participant was excluded for not following instructions on the working memory task. The final sample consisted of 37 academically stigmatized minorities (17 Hispanics, 16 African Americans, and 4 American Indian) and 40 Whites.

Procedure

Sessions were conducted with mixed-ethnic groups of 2 to 4 students by White experimenters. Upon entering the lab, participants initially heard an audio recording of a male researcher telling them that they would be completing several computerized tasks, including an initial verbal analogy test, other basic memory and categorization tasks, and finally a second verbal analogy test. All participants were told that the verbal analogy tests were highly predictive of performance on a variety of intelligence tests and that some people are better able to perform these tasks than others (instructions designed to elicit stereotype threat for minorities). Participants then completed demographic questions, including an item about their ethnicity.

To elicit anxiety about the later verbal analogy test, participants first completed five moderately difficult verbal analogy items that were taken from the GRE and were then asked to report the level of anxiety they were experiencing at the time.² Participants next completed a modified working memory task that contained sentences designed to prime either confidence or doubt (see the next section), followed by a questionnaire that included a manipulation check of stereotype threat.³ They were then informed that they would not be completing the second part of the verbal analogy test, were fully debriefed, thanked, and dismissed.

Materials

Anxiety. As a composite measure of anxiety ($\alpha = .92$), participants rated the extent to which they felt agitated, anxious, nervous, uneasy, and worried after completing the initial five verbal analogy items (1 = *not at all*, 7 = *extremely*).

Working memory task. The working memory task was adapted from the Reading Span Test (Daneman &

Carpenter, 1980). This computerized task is a dual-process measure in which participants alternate between memorizing words and reading sentences. Each trial began with the presentation of a neutral word on the middle of a computer screen for 2 s. Each word was then followed by a sentence that participants were instructed to read aloud. Participants were given 12 sets of these word-sentence combinations (sets had 4, 5, or 6 words for a total of 60 words). At the end of each set, participants were instructed to recall as many words as possible from the set they had just completed. Working memory was calculated using the absolute span scoring method (i.e., correctly recalled words were counted only for sets where recall was 100% correct; Turner & Engle, 1989).

Participants were randomly assigned to complete a version of the working memory task modified so that one sentence in each of the 12 sets contained either a confidence- or doubt-related word (e.g., "I am very *confident/doubtful* that our team will win the big game tonight"). Each sentence contained content that was irrelevant to intellectual achievement to minimize demand characteristics. In debriefing, no participant reported being aware of any deviation between the neutral sentences and the confidence- or doubt-related sentences embedded within the working memory task. To encourage participants to read the sentences, after each set, two sentences were shown on the screen and participants chose which sentence had actually been presented.

Manipulation check for stereotype threat. In the final questionnaire, participants rated their agreement on a 7-point scale with the two statements ($r = .87$, $p < .001$), "Thoughts of my ethnic group occurred frequently on my mind when I was working on the verbal IQ test" and "I am concerned that the researcher will judge people of my race/ethnicity based on my performance if I did not do well on today's IQ test." Responses to these two items were averaged to create an index of ethnicity-based evaluation concern.

Results

Manipulation Check of Stereotype Threat

To establish that the situation cued stereotype threat for minority participants, a 2 (prime: confidence vs. doubt) \times 2 (ethnicity: White vs. minority) ANOVA was conducted on our measure of ethnicity-based evaluation concerns. As predicted, minority students were more concerned about their performance reflecting on their group ($M_{\text{minority}} = 2.72$) than were Whites ($M_{\text{White}} = 1.41$), $F(1, 73) = 14.51$, $p < .001$.

Anxiety

To assess whether there were any condition differences in baseline anxiety, a 2 (prime: confidence vs. doubt) \times 2 (ethnicity: White vs. minority) ANOVA was performed on the self-reported anxiety scores. This analysis yielded no main effect for prime or a Prime \times Ethnicity interaction ($ps > .10$). There was, however, a marginal main effect for ethnicity, $F(1, 73) = 2.92$, $p = .09$. Minorities ($M = 3.28$, $SD = 1.55$, range = 1 to 7) reported marginally higher levels of anxiety compared to Whites ($M = 2.73$, $SD = 1.28$, range = 1 to 5.20). Thus, the stereotype threat manipulation engendered some feelings of anxiety among minority participants.

Working Memory

Sentence reading time. To ensure that our prime manipulation did not directly alter the way in which participants completed the working memory measure, we conducted an initial analysis of the average time participants took to read aloud sentences in the working memory task. A 2 (ethnicity: White vs. minority) \times 2 (prime condition: confidence vs. doubt) \times 2 (sentence type: prime vs. neutral) ANOVA (with sentence type as a within-subjects factor) yielded no main effects or interactions ($ps > .10$). Thus, participants read the neutral sentences at the same pace as the prime (i.e., confidence or doubt) sentences.

Word recall. Our primary hypothesis was that regardless of ethnicity, participants primed with doubt would show lower working memory to the extent that they reported higher levels of anxiety. When primed with confidence, the relationship between anxiety and working memory was expected to be attenuated if not reversed. To test this hypothesis, we conducted a series of hierarchical regression analyses (Aiken & West, 1991). The ethnicity of the participant, the type of prime they received, and the centered anxiety variable were entered in the first step of the model. The two-way interaction variables (e.g., Anxiety \times Prime) were entered in the second step of the model, and the three-way interaction variable was entered in the last step.

Results of this analysis yielded a main effect for anxiety ($\beta = -.25$, $p < .04$, $R^2 = .06$, all other $ps > .10$), replicating past research showing that anxiety is generally related to lower working memory (Klein & Boals, 2001). More important, this main effect was qualified by the predicted two-way interaction between anxiety and the type of prime ($\beta = -.35$, $p < .01$; see Figure 1). Simple slope analyses revealed that anxiety predicted lower working memory scores in the doubt condition ($\beta = -.73$, $p < .001$) but not in the confidence condition ($\beta = .02$, $p = .89$). Furthermore, at 1 *SD* above the

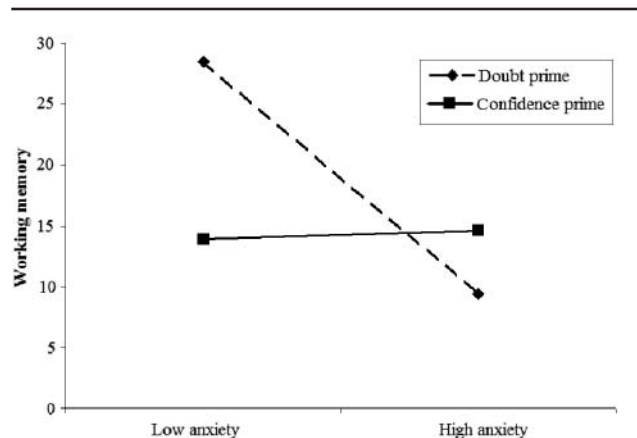


Figure 1 Study 1: Working memory scores (combined across ethnic group) as a function of anxiety and primed thoughts of confidence or doubt.

mean of anxiety, those primed with doubt had significantly lower working memory scores than those primed with confidence ($\beta = -.31$, $p < .05$). This pattern was reversed at 1 *SD* below the mean of anxiety ($\beta = .44$, $p < .01$).⁴ No other effects were significant ($ps > .24$), including the three-way interaction among ethnicity, prime type, and anxiety ($p = .62$). The lack of a prime main effect suggests that thoughts of doubt (vs. confidence) did not promote cognitive interference overall and that the lack of a significant three-way interaction with ethnicity reveals that the moderated pattern observed was not significantly different between White and minority participants.⁵

Discussion

Results from Study 1 supported our hypothesis that although individuals vary in the degree to which they experience anxiety in an intellectually threatening environment, the valence of their thoughts can alter whether that anxiety predicts cognitive impairments. When individuals experience heightened anxiety in an intellectual performance situation, they subsequently exhibited lower working memory but only when primed with thoughts of doubt. By comparison, anxiety did not predict lower working memory when participants had been primed with thoughts of confidence. This occurred regardless of the ethnicity of the participant, suggesting that metacognitive processes might play a general role in self-relevant performance situations, such as high-stakes intellectual testing. In other words, it has been argued that cues to group differences on a stereotyped ability might prime thoughts of doubt for stigmatized targets (Steele & Aronson, 1995) and thoughts of confidence for those who are positively stereotyped (Cheryan &

Bodenhausen, 2000). In the present study, we circumvented these default reactions by directly manipulating thoughts of confidence or doubt. As a result, Whites primed with doubt showed the same pattern that might be expected of minorities in an intellectually threatening situation. Similarly, minorities primed with confidence showed the same pattern that might be expected of Whites preparing to take a test of verbal ability.

Our approach in Study 1 was to hold constant the self-relevance of the performance situation (i.e., by describing the upcoming task as diagnostic of intelligence) to demonstrate that a subtle manipulation of accessible thoughts could moderate the relationship between anxiety and working memory among both stigmatized and nonstigmatized individuals. However, primed thoughts of doubt (or confidence) are unlikely to have the same effect in situations that are less self-relevant. Research examining the metacognitive processes that influence persuasion, for example, have found these effects only among individuals motivated to engage in deliberative thought (Briñol, Petty, & Barden, 2007), which is often manipulated by varying the self-relevance of the situation. In other words, one must be motivated to think about the current context to weigh currently activated thoughts and emotions as evidence for one's opinion.

Similarly, in performance situations, Tesser (1988) has argued that self-evaluation maintenance processes are most likely to be triggered during tasks that are high in self-relevance, such as a diagnostic test. Thus, just as the self-relevance of a domain is a key variable that prompts defensive self-esteem repair in the face of threatening social comparisons, we suggest that the metacognitive processes we tested in Study 1 are only likely to be prompted in situations framed as being self-relevant because this is when people will be motivated to evaluate themselves online during their performance. When a task is irrelevant to one's sense of self-integrity, there should be little motivation to engage in performance-monitoring processes to assess how one is doing (Forbes, Schmader, & Allen, 2008). In fact, a common way to experimentally eliminate stereotype threat is to simply describe a task in a more neutral way, as not diagnostic of any particular ability (e.g., Steele & Aronson, 1995). Thus, to test our assumption that the metacognitive process we have proposed only occurs under conditions of high self-relevance, Study 2 was designed to examine the interaction between anxiety and thought valence among a sample of women preparing to complete a task described as either a diagnostic math test or a neutral problem-solving activity.

We hypothesized that when the situation was described as a diagnostic test, women primed with thoughts of doubt would show a negative relationship

between their anxiety levels and working memory. However, priming women with thoughts of confidence was expected to attenuate if not reverse this pattern. In contrast, when the primary task was described as a simple problem-solving exercise, we expected our manipulation of confidence and doubt to have little effect on working memory, either directly or in interaction with initial anxiety levels.

STUDY 2

Method

Participants and Design

Participants were 116 female undergraduates who self-identified as American. All participants received course credit for their time. Three participants were excluded for not following instructions on the working memory task. Two additional participants with working memory scores over 3 *SD* above the mean were also excluded. The final sample consisted of 111 females (79 White, 10 Hispanic, 7 African American, 7 Asian American, 1 American Indian, and 7 unidentified) randomly assigned to condition in a 2 (task frame: diagnostic vs. neutral) \times 2 (prime: confidence vs. doubt) between-subjects design.

Procedure

Each session was conducted with 2 to 4 participants by a White male and female experimenter. The procedure was the same as Study 1 except that the central task in this study was either described as a math test or a problem-solving task (Johns et al., 2005). Women in the diagnostic condition were told that the two-part math test they were going to take was highly predictive of people's natural mathematical ability and that we were interested in comparing men's and women's performance on the diagnostic math test to help us establish norms for the two groups. Females in the neutral condition were informed that the two-part task they were going to complete was merely a problem-solving exercise and that we were interested in comparing individual performances. Participants then completed demographic questions, including gender in the diagnostic condition only.

Parallel to Study 1, participants then completed five difficult math items taken from a practice GRE exam and then completed the same anxiety measure used in Study 1 ($\alpha = .80$). They then performed the same working memory measure modified to prime confidence or doubt as well as a final questionnaire including a modification of the two manipulation check items measured

in Study 1 ($r = .70, p < .001$), which were then averaged to reflect women's concerns that their performance would be used to judge women as a whole.

Results

Manipulation Check of Stereotype Threat

A 2 (prime: confidence vs. doubt) \times 2 (task frame: diagnostic vs. neutral) ANOVA was conducted on participants' ratings of concern that the researcher would judge women's abilities based on their personal performance. As predicted, women led to believe that they would be taking a diagnostic math test reported more concern about how their gender would be evaluated ($M = 2.08$) compared with women in the neutral condition ($M = 1.52$), $F(1, 103) = 6.24, p < .05$.⁶ No other effects were significant ($ps > .10$). Thus, our manipulation of stereotype threat was effective in eliciting group-based performance concerns.

Anxiety Ratings

A 2 (prime: confidence vs. doubt) \times 2 (task frame: diagnostic vs. neutral) ANOVA was conducted on self-reported anxiety ratings taken after the initial math/problem-solving items. This analysis produced an unexpected interaction, $F(1, 107) = 3.83, p = .05$. Although the simple effects tests were not significant, women subsequently assigned to the doubt prime condition reported somewhat more anxiety in the diagnostic condition ($M = 2.74$) compared to the neutral condition ($M = 2.34$), $p = .21$. However, anxiety ratings were reversed between the diagnostic ($M = 2.50$) and neutral ($M = 2.93$) conditions among women subsequently assigned to receive confidence primes, $p = .13$. As our primary hypothesis concerns the relationship that these anxiety ratings have with working memory, rather than simple mean differences in working memory between the four conditions, these initial condition differences in anxiety are partialled out of our analysis of working memory, minimizing any effect they could have.

Working Memory

Time spent reading sentences. As in Study 1, we first checked to ensure that participants did not spend different amounts of time reading prime or neutral sentences. A 2 (task frame) \times 2 (prime) \times 2 (sentence type) ANOVA (with sentence type as a within-subjects factor) conducted on the average time participants spent reading aloud confidence/doubt versus neutral sentences in the working memory task revealed no significant main effects or interactions ($ps > .05$).

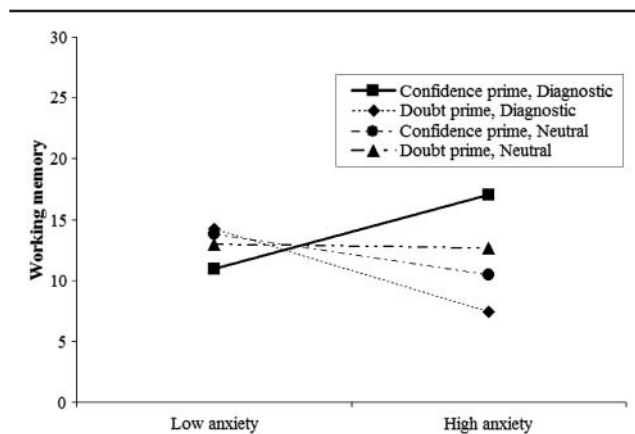


Figure 2 Study 2: Working memory scores as a function of task frame, anxiety, and primed thoughts of confidence or doubt.

Word recall. We again conducted a hierarchical regression analysis to test our primary hypothesis that only in a stereotype-threatening situation would priming women with thoughts of doubt (compared with confidence) produce a negative relationship between initial anxiety and working memory. Thus, working memory scores were regressed on task frame, prime, and the centered anxiety variable on Step 1, the two-way interactions between these predictors on Step 2, and the three-way interaction variable on Step 3.

This analysis revealed no significant main effects or two-way interactions ($ps > .25$); however, the predicted three-way interaction among task frame, prime, and anxiety level was significant, ($\beta = -.20, p = .05$; see Figure 2). Simple interaction analyses revealed that the Anxiety \times Prime interaction observed in Study 1 was only significant in the diagnostic task frame condition ($\beta = -.30, p < .04$) and was not significant in the neutral task frame condition ($\beta = .10, p > .10$). Simple slope analyses within the diagnostic task frame condition further revealed that at 1 SD above the mean of anxiety, women primed with doubt had significantly lower working memory scores compared with women primed with confidence ($\beta = -.45, p < .03$). At 1 SD below the mean of anxiety, the prime manipulation had no effect on working memory ($\beta = .15, p = .42$). In addition, within the diagnostic task frame condition, women primed with doubt who reported higher levels of anxiety tended to have lower working memory scores compared with women who reported lower levels of anxiety ($\beta = -.35, p = .08$). This pattern was reversed but was not significant among women who received confidence primes ($\beta = .26, p = .22$). Thus, consistent with our hypothesis, primed thoughts of doubt significantly altered the relationship between anxiety and working memory, but only among women expecting to take a diagnostic math test.

Discussion

Results of Study 2 replicated findings in Study 1 with a different stigmatized group. Primed thoughts of confidence or doubt moderated the relationships that anxiety had with women's working memory only when women were placed in a stereotype-threatening environment. Specifically, women who experienced higher levels of anxiety under stereotype threat exhibited somewhat lower working memory when they were primed with doubt. Conversely, priming women with confidence buffered them from this relationship under stereotype threat. However, when the situation was framed in a less self-relevant way, anxiety did not predict women's working memory. Presumably, this less threatening frame did not promote the level of self-evaluation required to elicit metacognitive interpretation of anxiety in light of primed thoughts of doubt.

STUDY 3

Whereas the previous two studies tested our metacognitive hypothesis by experimentally manipulating the thoughts that people have within a performance context, Study 3 was designed to conceptually replicate these effects in a more naturalistic setting using an individual difference approach. Specifically, if situations of stereotype threat typically induce negative interpretations of arousal but priming thoughts of confidence eliminates this relationship, individuals who are dispositionally inclined to reappraise negative feelings in a more positive way should not show a negative relationship between arousal and performance. In fact, for reappraisers, greater arousal should predict better performance (e.g., Blascovich, Mendes, Hunter, & Salomon, 1999). Thus, in Study 3, we sought to test reappraisal tendencies as a moderator of the relationship between arousal and performance.

In addition, Study 3 included several other changes designed to address limitations in the prior experiments. First, although we argue that metacognitive processes should be spontaneously cued for stigmatized individuals when a task is both self-relevant and stereotypical, Study 2 confounded these two factors. In Study 3, we chose a highly relevant performance situation, namely, women taking a GRE practice test in preparation for the actual exam. Although both the math and verbal subtests should be highly self-relevant to graduating seniors, the metacognitive processes we discuss here should be most likely to operate when women perform under the burden of a negative stereotype (i.e., during the math subtest but not during the verbal subtest). Furthermore,

examining these processes in this context allows us to directly assess women's performance on math and verbal problems rather than working memory processes that are thought to underlie performance impairments due to threat.

Finally, whereas Studies 1 and 2 looked at one specific type of arousal by measuring self-reported anxiety, Study 3 measured arousal physiologically. Although even discrete emotions can be subject to further metacognitive framing (e.g., worrying about worrying), we expect these metacognitive processes to also play a role in one's initial interpretation of an undifferentiated state of physiological arousal. In Study 3, we operationalized arousal as an end-product of one of the two primary stress systems, specifically, the sympathetic adrenal medullary system. Although this system is often measured by electrophysiological changes in the autonomic nervous system, we opted for a noninvasive approach that allowed us to examine sympathetic activation detected in saliva. Recently, salivary alpha-amylase (sAA) activity has been used as an estimate of increases in catecholamines (specifically, epinephrine and norepinephrine) associated with increased sympathetic nervous system activity in response to psychosocial stressors (e.g., Rohleder, Wolf, Maldonado, & Kirschbaum, 2006). In addition to regulating glucocorticoids in the brain, catecholamines also are influenced by changes in alertness and affect (see Lovallo & Thomas, 2000). As increases in sympathetic activation occur within seconds of exposure to a stressor, and peak responses occur within 5 min, we obtained saliva samples just before women began both the math and verbal subtests under the assumption that sAA levels would peak at this point among those most aroused (Granger et al., 2006).

In this study, we assert that individual differences in the tendency to reappraise negative affect should interact with sympathetic activation to predict women's performance on a math subtest (but not a verbal subtest). Specifically, for women who experience heightened activation when taking a practice GRE test, those who are low in reappraisal tendencies should perform more poorly on the math problems and report more self-doubt about their math performance than those who are high in reappraisal tendencies. For those who do not experience heightened activation in the first place, reappraisal should have little relationship to math performance. Furthermore, because the verbal subtest does not have implications for gender stereotypes (Spencer et al., 1999), these same interpretative processes should not be predictive of performance in that context.

Method

Participants

Participants were 43 English-speaking female undergraduates who were preparing to take the GRE (M age = 23, SD = 4; 79% Caucasian). All were screened for depression, anxiety, malaria, mononucleosis, or a recent surgical procedure that could affect neuroendocrine concentrations. Participants refrained from smoking, physical exercise, and food for 2 hr before their participation and were compensated with \$15 and a GRE practice guide.⁷

Procedure

All participants arrived at the lab at 2 p.m. for a study examining the role of arousal and test performance. To assess baseline sAA on a nontest day, participants provided a saliva sample during an initial survey session that occurred on a different day (typically 1-3 days before the main session). In both sessions, participants first rested for 20 min while they completed surveys. In the main session, participants were then placed in individual cubicles to take computerized versions of the math and verbal GRE practice tests. Order was determined randomly and participants had 30 min to work on each subtest. They provided a saliva sample before they were assigned each subtest. Thus, sympathetic activation was assessed before they knew which test they were about to take. They completed a final questionnaire at the end of the session. Before being debriefed, participants received their practice test scores in a sealed envelope.

Measures

Performance. The verbal and the math subtests were presented in MediaLab and consisted of 38 questions and 30 questions, respectively, taken from *Practice to Take the GRE General Test* (10th ed.). Participants were required to answer each question (presented in a fixed order) before proceeding to the next, and previously completed questions could not be reviewed.

Reappraisal tendency. Before taking the tests, participants completed Gross and John's (2003) six-item reappraisal scale (α = .89; e.g., "When I'm faced with a stressful situation, I make myself think about it in a way that helps me stay calm"; 1 = *strongly disagree*, 7 = *strongly agree*).

Self-doubt. In the final questionnaire, participants rated their feelings of self-doubt experienced while taking each practice test with six items: "doubts about my ability," "foolish," "inferior to others," "insecure about

my performance," "unsure of myself," "confident about how I was doing"—reverse coded (1 = *not at all*, 7 = *extremely*; α s = .84 for quantitative section and .86 for verbal section).

sAA. Saliva samples (.5 to 1 mL each) were obtained with the spitting method (Navazesh, 1993). Participants were instructed to first void their mouths of saliva by swallowing. They then had 5 min to expectorate saliva via a straw into a sterile tube, which was then immediately sealed, coded by subject number, and stored at -20°C until shipped to a laboratory in Dresden, Germany, where it was thawed and then centrifuged at 3,000 rpm for 5 min. Concentration of alpha-amylase in saliva was measured by an enzyme kinetic method (α -amylase EPS Sys; Roche Diagnostics, Mannheim, Germany; see Rohleder et al., 2006, for additional details).

Results and Discussion

Manipulation Checks for Stereotype Threat and Self-Relevance

The nontest-day survey included a three-item measure of the self-relevance of performance on the math and verbal GRE tests adapted from Major and Schmader's (1998) disengagement scale (e.g., "It really doesn't matter to me one way to the other how I do on the math [verbal] section of the GRE"—reverse scored). As expected, women reported that both the verbal GRE (M = 5.17, on a 1-7 scale) and the math GRE (M = 4.96) were equivalently highly self-relevant, $t(42) = 1.51$, $p > .10$.

The final questionnaire asked participants to rate the researcher's expectations for men's versus women's performance on both the math and verbal subtests (1 = *men do better*, 7 = *women do better*). As hypothesized, women thought that the researcher expected women to perform better than men on the verbal (M = 4.81) as compared with the math test (M = 3.07), $t(42) = -5.97$, $p < .001$. Thus, consistent with our assumptions, women found both tests to be highly self-relevant but perceived stereotype threat only on the math subtest.

Test Performance

To control for individual variation in sAA concentrations, sAA assessments taken before both the math and verbal subtests were residualized for sAA levels taken on the nontest day. In the interest of parsimony, these residualized versions of sAA will simply be referred to as math sAA and verbal sAA. To test our core hypothesis, we regressed math performance on our centered reappraisal and math sAA variables (Step 1) and their interaction (Step 2). Reappraisal tendencies were not associated with math ($r = -.10$) and verbal ($r = -.14$) sAA, indicating that

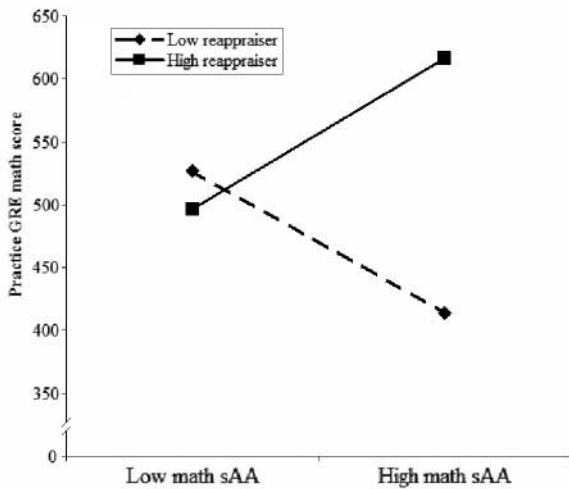


Figure 3 Study 3: Practice Graduate Record Exam (GRE) math scores as a function of reappraisal tendency and math salivary alpha-amylase (sAA).

multicollinearity is not a problem in these analyses. Results yielded only the predicted interaction ($\beta = .63$, $p < .01$; Figure 3). Simple slopes analyses revealed that at high levels of math sAA (+1 SD), reappraisal predicted better performance ($\beta = .90$, $p < .001$), but at low levels (−1 SD) this relationship was not significant ($\beta = -.36$, $p > .05$). In addition, among low reappraisers (−1 SD), math sAA predicted poorer performance ($\beta = -.62$, $p < .05$), whereas this relationship was actually positive for high reappraisers (+1 SD; $\beta = .65$, $p < .01$).

When the same analysis was conducted for verbal test performance (using verbal sAA as a predictor), no effects, including the interaction ($\beta = -.15$), were significant (all $ps > .05$). Thus, as predicted, initial arousal predicted poorer performance only on a stereotype-relevant task and only among nonreappraisers. In contrast, reappraisers performed better on the math test to the degree that their arousal level was high.

Self-Reported Doubt

We also tested whether reappraisal tendencies interact with sympathetic activation to predict women's perceptions of self-doubt about their math and verbal performance. Using the same moderated regression approach, we found a significant interaction between reappraisal and sAA for ratings of math doubt ($\beta = -.45$, $p < .05$; Figure 4). At high math sAA, reappraisal predicted lower math doubt ($\beta = -.52$, $p = .05$), but this relationship was not significant at low math sAA ($\beta = .39$, $p > .05$). Also, among high reappraisers (+1 SD), math sAA predicted lower math doubt ($\beta = -.64$, $p < .01$), whereas this relationship was not significant for low reappraisers (−1 SD; $\beta = .26$, $p > .05$).

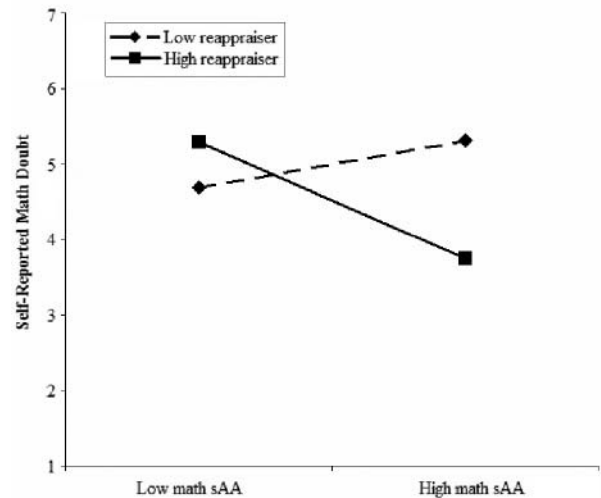


Figure 4 Self-reported math doubt during math Graduate Record Exam (GRE) subtest as a function of reappraisal tendency and math salivary alpha-amylase (sAA).

The same analysis on doubt about one's verbal performance revealed no effects (all $ps > .10$). Again, only in a situation of stereotype threat did reappraisal tendencies moderate how sympathetic activation relates to one's interpretation of their performance.

GENERAL DISCUSSION

The present set of studies draws from an integrated process model of stereotype threat proposed by Schmader et al. (2008) to test how affective and cognitive variables that have been previously studied in isolation interact to predict impairments to working memory and intellectual test performance. In Studies 1 and 2, we hypothesized that subtle primes of doubt or confidence would moderate the relationship that initial anxiety has on working memory within an intellectually threatening performance context. In Study 1, anxiety was related to lower working memory for participants primed with thoughts of self-doubt but not for participants primed with thoughts of self-confidence. Interestingly, this same pattern of findings was observed among minority and White participants, suggesting that group differences in cognitive performance observed in past studies (e.g., Schmader & Johns, 2003) might be the result of stereotype threat triggering relatively more negative self-thoughts spontaneously for minority college students compared to their White counterparts (Steele & Aronson, 1995).

In Study 2, women exposed to stereotype threat showed lower working memory to the degree that they reported high anxiety *and* had been primed with doubt as compared to confidence. Interestingly, anxiety did

not interact with primed thoughts to predict working memory when the study was described in neutral terms, allowing us to rule out the possibility that simply pairing anxiety with negative thoughts will necessarily impair cognitive processing. Rather, these results are consistent with the idea that negatively stereotyped individuals are especially likely to engage in metacognitive monitoring processes when the situation is highly self-relevant (Schmader et al., 2008).

Prior studies have shown that negative thoughts are spontaneously activated by stereotype threat (Beilock et al., 2007; Cadinu et al., 2005). However, because these studies measured, rather than manipulated, negative thoughts, it is possible that anxiety would have also been increased by the same threat manipulations leading to inherent confounds between these two constructs. By directly manipulating thought valence orthogonal to anxiety, we have extended our understanding of the interplay between these cognitive and affective mechanisms. If anxiety and negative thinking are independent predictors of cognitive deficits due to stereotype threat, we would have expected these variables to show only main effects on performance: lower performance when anxious and lower performance when primed with doubt. However, the thought prime manipulation in Studies 1 and 2 did not have a main effect on working memory, and anxiety only had a main effect relationship with working memory in Study 1. The pattern of results that replicated across both studies thus provides stronger support for a model where affect and cognitive variables interact with one another in predicting cognitive impairments known to underlie stereotype threat effects on performance.

Having shown the interactive effects of affect and primed thoughts on cognitive efficiency, Study 3 then tested these metacognitive processes with a different set of variables in a more naturalistic testing context and established that these effects happen only when negative stereotypes apply. Specifically, we showed that among women who approached a mock GRE exam with higher levels of sympathetic activation, those who typically reappraise negative emotions performed better on the math subtest relative to women low in reappraisal tendencies. When activation was low, reappraisal was not linked to math performance. Self-reported experiences of doubt provided further evidence that reappraisal tendencies were associated with differences in the cognitive interpretation of arousal. When sympathetic activation was high, women who tend to reappraise negative emotions experienced less doubt about their math ability than those low in reappraisal. Again, at low levels of activation, reappraisal and doubt were unrelated. This final study highlights the importance of the cognitive interpretation of arousal as a powerful and reliable

predictor of cognitive performance. Furthermore, the fact that these effects were not observed on verbal test performance is consistent with our contention that metacognitive processes play a distinct role under stereotype threat and do not necessarily operate in any self-relevant performance situation.

It is also noteworthy that the predicted metacognitive processes were apparent when arousal was operationalized in terms of either a discrete negative emotion (i.e., anxiety) or an undifferentiated physiological state (i.e., sympathetic activation). Both kinds of affective response have been studied in prior stereotype threat research under the assumption that they could explain lower performance. Here we suggest that the relationship is moderated by other cognitions activated by the situation or an individual's coping style. Consistent with the two-factor theory of emotion (Schacter & Singer, 1962), sympathetic activation means something different if it is appraised in a positive way. But such appraisal processes do not only moderate the meaning of physiological activation, they can also moderate the meaning of felt emotions themselves. Although an emotion such as anxiety typically has negative connotations, results of Studies 1 and 2 show that anxiety ceases to predict lower cognitive resources if it is experienced alongside confidence primes.

The Role of Positive Framing in Stereotype Threat

The results of the present studies highlight the role of having a positive mindset in moderating susceptibility to stereotype threat. For example, the results of Study 3 are consistent with other research showing that people's interpretation of their affective states can moderate stereotype threat effects (Ben-Zeev et al., 2005; Johns et al., 2005; Johns et al., 2008; Stone, Lynch, Sjomeling, & Darley, 1999; Walton & Cohen, 2007). But these past studies have always manipulated such interpretations directly and only with the intent of deflecting arousal from the self rather than providing a positive interpretation of it. Study 3 shows that individual differences in the tendency to positively appraise one's affective states actually predicts better performance in a stereotype-relevant task if a person is experiencing high sympathetic activation. Positive reappraisal can do more than just reduce the effects of threat, it can actually fuel better performance. Thus, this last study sheds light on the importance of differentiating "good" from "bad" arousal. Not all stress responses are created equal and though increases in sympathetic nervous system activation (SNS) often are interpreted as having a negative connotation, this is not necessarily the case. Increases in SNS can also predict better cognitive performance and positive affect (Mendes, Major, McCoy, & Blascovich, 2008).

Further benefits of positive framing were found in the first two studies where priming individuals with thoughts of confidence seemed to circumvent the cycle whereby anxiety predicts cognitive impairments. It is possible that manipulations such as providing a positive role model (Marx & Roman, 2002) or self-affirmation (Martens, Johns, Greenberg, & Schimel, 2006) are effective at reducing stereotype threat because they prime self-confidence. Likewise, individuals who are able to excel in domains in which they are negatively stereotyped might also be able to resist the metacognitive processes that normally impair executive functioning by maintaining an unshakable sense of confidence that is impervious to the activation of negative stereotypes.

On the surface, this suggestion seems to contradict Steele's (1997) vanguard hypothesis, which suggests that those who excel in a domain would be most susceptible to experience stereotype threat. However, there is an important conceptual distinction between believing in one's ability to succeed and being highly invested in good performance. Investment, often operationalized as domain identification or psychological engagement, has been shown to increase susceptibility to stereotype threat (Aronson et al., 1999; Stone et al., 1999), but this does not necessarily mean that perceptions of competence would have the same effect. *Caring* that one can do well might increase susceptibility to stereotype threat at the same time that *believing* that one can do well might reduce it. Future research is needed to explore this possibility.

Conclusions

Over the past two decades, social psychological research has placed a premium on understanding the psychological mechanisms that underlie basic situational effects. However, complex social phenomena such as stereotype threat are unlikely to be explained by a single underlying process (Schmader et al., 2008; Steele et al., 2002), and research is sorely needed to examine the complex interplay of distinct mechanisms. The present studies take this integrative approach to show that anxious arousal and negative thinking do not predict cognitive impairments in isolation. Rather, affective and cognitive variables combine in a metacognitive way such that performance is impaired in high-stakes situations where one's initial affective experience is interpreted in light of concurrent negative thoughts. As this combination of processes is more likely to be experienced by those who are negatively stereotyped, the present results are supportive of an integrated process model of stereotype threat whereby a motivation to monitor one's performance at a self-relevant task is biased by the activation of negative group stereotypes. However, this biased processing can be circumvented by situations that prime positive thoughts or for

individuals with a tendency to reappraise their negative feelings. Establishing the interplay between cognitive and affective variables not only extends our understanding of stereotype threat, it also sheds light on the individual difference variables and situational manipulations that reduce one's susceptibility to the pernicious effects that negative stereotypes can have on performance.

NOTES

1. Participants' knowledge of intelligence-related stereotypes was assessed in a pretest. White participants were included in the study if they indicated that society perceived Whites to have higher intelligence than Blacks, Latinos, or American Indians. Minority participants were included if they indicated that society perceived Whites to have higher intelligence than their ethnic group specifically.

2. Performance on these initial items was not of interest because five questions would not produce enough variability in responding.

3. Studies 1 and 2 included exploratory measures of implicit and explicit domain identification, performance expectancies, and state self-esteem that did not yield meaningful results and thus will not be discussed further, with one exception footnoted in the results.

4. It is not clear why individuals who reported particularly low levels of anxiety showed significantly higher working memory when primed with doubt as compared to confidence. Perhaps in the absence of anxiety, such individuals are able to successfully react against primed negative thoughts. However, as this unpredicted pattern does not replicate in Study 2, it would be premature to draw firm conclusions from it.

5. Additional analyses revealed an Anxiety \times Prime interaction on a measure of participants' performance expectations ($\beta = -.31$, $p < .05$). Regardless of ethnicity, anxiety predicted more negative expectancies when participants were primed with doubt ($\beta = -.45$, $p < .05$) but was unrelated to expectancies among participants primed with confidence ($\beta = .19$, *ns*). This same pattern was evident in Study 2 but was not significant ($\beta = -.12$, $p > .10$).

6. Degrees of freedom are lower on this variable because 4 participants were unable to complete the final questionnaire in the allotted time.

7. Data reported in Study 3 were part of a larger study including additional measures not relevant to these hypotheses. Although 21 men were also examined, this male sample was not large enough to provide a valid test of these hypotheses. The larger study also included a reappraisal manipulation. However, as initial analyses suggested that this manipulation was ineffective, analyses collapse across this manipulation but control for it.

REFERENCES

- Aiken, L. S., & West, S. G. (1991). *Multiple regression: Testing and interpreting interactions*. Thousand Oaks, CA: Sage.
- 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.
- Beilock, S. L., Rydell, R. J., & McConnell, A. R., (2007). Stereotype threat and working memory: Mechanisms, alleviation, and spill-over. *Journal of Experimental Psychology: General*, 136, 256-276.
- Ben-Zeev, T., Fein, S., & Inzlicht, M. (2005). Arousal and stereotype threat. *Journal of Experimental Social Psychology*, 41, 174-181.
- Blascovich, J., Mendes, W. B., Hunter, S., & Salomon, K. (1999). Social facilitation as challenge and threat. *Journal of Personality and Social Psychology*, 77, 68-77.
- 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.
- Bosson, J. K., Haymovitz, E. L., & Pinel, E. C. (2004). When saying and doing diverge: The effects of stereotype threat on self-reported

- versus non-verbal anxiety. *Journal of Experimental Social Psychology*, 40, 247-255.
- Briñol, P., & Petty, R. E. (2003). Overt head movements and persuasion: A self-validation analysis. *Journal of Personality and Social Psychology*, 84, 1123-1139.
- Briñol, P., Petty, R. E., & Barden, J. (2007). Happiness versus sadness as a determinant of thought confidence in persuasion: A self-validation analysis. *Journal of Personality and Social Psychology*, 93, 711-727.
- Cadinu, M., Maass, A., Rosabianca, A., & Kiesner, J. (2005). Why do women underperform under stereotype threat? Evidence for the role of negative thinking. *Psychological Science*, 16, 572-578.
- Cheryan, S., & Bodenhausen, G. V. (2000). When positive stereotypes threaten intellectual performance: The psychological hazards of "model minority" status. *Psychological Science*, 11, 399-402.
- Croizet, J. C., Despres, G., Gauzins, M., Huguot, P., & Leyens, J. (2004). Stereotype threat undermines performance by triggering a disruptive mental load. *Personality and Social Psychology Bulletin*, 30, 721-731.
- Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior*, 19, 450-466.
- Educational Testing Service. (2002). *Practice to take the GRE General Test* (10th Ed.). Princeton, NJ: Educational Testing Service.
- Forbes, C. E., Schmader, T., & Allen, J. J. B. (2008). Error monitoring in an intellectually threatening environment. *Social Cognitive Affective Neuroscience*, 3, 253-261.
- Granger, D. A., Kivlighan, K. T., Blair, C., El-Sheikh, M., Mize, J., Lisonbee, J. A., et al. (2006). Integrating the measurement of salivary alpha amylase into studies of child health, development, and social relationships. *Journal of Personal and Social Relationships*, 23, 267-290.
- Inzlicht, M., McKay, L., & Aronson, J. (2006). Stigma as ego depletion: How being the target of prejudice affects self-control. *Psychological Science*, 17, 262-269.
- Johns, M. J., Inzlicht, M., & Schmader, T. (2008). Stereotype threat and executive resource depletion: Examining the influence of emotion regulation. *Journal of Experimental Psychology: General*, 137, 691-705.
- Johns, M., & Schmader, T. (in press). Metacognitive regulation of feeling states in stereotype threat. In R. M. Arkin, K. C. Oleson, & P. J. Carroll (Eds.), *The uncertain self: A handbook of perspectives from social and personality psychology*. Mahwah, NJ: Lawrence Erlbaum.
- Johns, M., Schmader, T., & Martens, A. (2005). Knowing is half the battle: Teaching stereotype threat as a means of improving women's math performance. *Psychological Science*, 16, 175-179.
- Jost, J. T., Kruglanski, A. W., & Nelson, T. O. (1998). Social metacognition: An expansionist review. *Personality and Social Psychology Review*, 2, 137-154.
- Klein, K., & Boals, A. (2001). The relationship of life event stress and working memory capacity. *Applied Cognitive Psychology*, 15, 565-579.
- Lazarus, R. S. (1991). Cognition and motivation in emotion. *American Psychologist*, 46, 352-367.
- Lovaglio, W. R., & Thomas, T. L. (2000). Stress hormones in psychophysiological research: Emotional, behavioral, and cognitive implications. In J. T. Cacioppo, L. G. Tassinary, & G. G. Berntson (Eds.), *Handbook of psychophysiology* (2nd ed., pp. 342-367). Cambridge, UK: Cambridge University Press.
- Major, B., & Schmader, T. (1998). Coping with stigma through psychological disengagement. In J. Swim & C. Stangor (Eds.), *Prejudice: The target's perspective* (pp. 220-243). San Diego, CA: Academic Press.
- Martens, A., Johns, M., Greenberg, J., & Schimel, J. (2006). Combating stereotype threat: The effect of self-affirmation on women's intellectual performance. *Journal of Experimental Social Psychology*, 46, 236-243.
- Marx, D. M., & Roman, J. S. (2002). Female role models: Protecting female students' math test performance. *Personality and Social Psychology Bulletin*, 28, 1185-1197.
- Marx, D. M., & Stapel, D. A. (2006). It's all in the timing: Measuring emotional reactions to stereotype threat before and after taking a test. *European Journal of Social Psychology*, 36, 687-698.
- Mendes, W. B., Blascovich, J., Lickel, B., & Hunter, S. (2002). Challenge and threat during social interaction with White and Black men. *Personality and Social Psychology Bulletin*, 28, 939-952.
- Mendes, W. B., Major, B., McCoy, S., & Blascovich, J. (2008). How attributional ambiguity shapes physiological and emotional responses to social rejection and acceptance. *Journal of Personality and Social Psychology*, 94, 278-291.
- Murphy, M., Steele, C., & Gross, J. (2007). Signaling threat: How situational cues affect women in math, science, and engineering. *Psychological Science*, 18, 879-885.
- Navazesh, M. (1993). Methods for collecting saliva. *Annals of the New York Academy of Science*, 694, 72-77.
- Rohleder, N., Wolf, J. M., Maldonado, E. F., & Kirschbaum, C. (2006). The psychosocial stress-induced increase in salivary alpha-amylase is independent of saliva flow rate. *Psychophysiology*, 43, 645-652.
- Schachter, S., & Singer, J. E. (1962). Cognitive, social and physiological determinants of emotional states. *Psychological Review*, 69, 379-399.
- 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., & Forbes, C. (2008). An integrated process model of stereotype threat effects on performance. *Psychological Review*, 115, 336-356.
- Schwarz, N., & Clore, G. L. (1983). Mood, misattribution and judgments of well-being: Informative and directive functions of affective states. *Journal of Personality and Social Psychology*, 45, 513-523.
- Shih, M., Ambady, N., Richeson, J. A., Fujita, K., & Gray, H. M. (2002). Stereotype performance boosts: The impact of self-relevance and the manner of stereotype activation. *Journal of Personality and Social Psychology*, 83, 638-647.
- 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.
- Stangor, C., Carr, C., & Kiang, L. (1998). Activating stereotypes undermines task performance expectations. *Journal of Personality and Social Psychology*, 75, 1191-1197.
- Steele, C. M. (1997). A threat in the air: How stereotypes shape the intellectual identities and performance. *American Psychologist*, 52, 613-629.
- 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.
- 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). New York: 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.
- Tesser, A. (1988). Toward a self-evaluation maintenance model of social behavior. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 21, pp. 181-227). New York: Academic Press.
- Turner, M. L., & Engle, R. W. (1989). Is working memory task dependent? *Journal of Memory and Language*, 28, 127-154.
- Walton, G. M., & Cohen, G. L. (2007). A question of belonging: Race, social fit, and achievement. *Journal of Personality and Social Psychology*, 92, 82-96.
- Watkins, E. (2004). Appraisals and strategies associated with rumination and worry. *Personality and Individual Differences*, 37, 679-694.
- Wells, A. (1995). Metacognition and worry: A cognitive model of generalised anxiety disorder. *Behavioural and Cognitive Psychotherapy*, 23, 301-320.
- Wheeler, S. C., & Petty, R. E. (2001). The effects of stereotype activation on behavior: A review of possible mechanisms. *Psychological Bulletin*, 127, 797-826.

Received May 15, 2008

Revision accepted November 12, 2008