
On the Causal Mechanisms of Stereotype Threat: Can Skills That Don't Rely Heavily on Working Memory Still Be Threatened?

Sian L. Beilock

The University of Chicago

William A. Jellison

Colgate University

Robert J. Rydell

University of California, Santa Barbara

Allen R. McConnell

Miami University

Thomas H. Carr

Vanderbilt University

Recent work suggests that stereotype threat (ST) harms performance by reducing available working memory capacity. Is this the only mechanism by which ST can occur? Three experiments examined ST's impact on expert golf putting, which is not harmed when working memory is reduced but is hurt when attention is allocated to proceduralized processes that normally run outside working memory. Experiment 1 showed that well-learned golf putting is susceptible to ST. Experiments 2 and 3 demonstrated that giving expert golfers a secondary task eliminates ST-induced impairment. Distracting attention away from the stereotype-related behavior eliminates the harmful impact of negative stereotype activation. These results are consistent with explicit monitoring theories of choking under pressure, which suggest that performance degradation can occur when too much attention is allocated to processes that usually run more automatically. Thus, ST alters information processing in multiple ways, inducing performance decrements for different reasons in different tasks.

Keywords: *stereotype threat; automaticity; expertise; choking; performance pressure*

Introducing a negative stereotype about a social group in a particular domain can reduce the quality of performance exhibited by members of that group (Steele, 1997).

For example, when negative stereotypes are activated, African Americans perform worse on tasks described as assessing intelligence (e.g., Steele & Aronson, 1995), Whites perform worse on tasks described as assessing natural athletic ability (e.g., Stone, Lynch, Sjomeling, & Darley, 1999), and women perform worse on math-related tasks (e.g., Spencer, Steele, & Quinn, 1999). Although the prevalence of stereotype threat effects has been widely demonstrated across many diverse social groups and task types, relatively less is known about the cognitive processes that underlie these effects (Wheeler & Petty, 2001). That is, how does activating a negative performance-related stereotype lead to less-than-optimal skill execution among members of the stereotyped group?

Authors' Note: This research was supported by Institute of Educational Sciences (IES) Grant R305H050004 to Sian Beilock and National Science Foundation (NSF) Grant BCS-0601148 to Sian Beilock and Allen McConnell. Please address correspondence to Sian L. Beilock, Department of Psychology, 5848 South University Avenue, University of Chicago, Chicago, IL 60637; e-mail: beilock@uchicago.edu.

PSPB, Vol. 32 No. 8, August 2006 1059-1071

DOI: 10.1177/0146167206288489

© 2006 by the Society for Personality and Social Psychology, Inc.

An understanding of how stereotype threat exerts its impact will not only contribute to our theoretical understanding of this important phenomenon but will also shed light on the types of training and performance strategies that may help alleviate these unwanted skill decrements (e.g., Johns, Schmader, & Martens, 2005). Moreover, because stereotype threat has been demonstrated across diverse task domains, ranging from math problem solving to golf putting—skills known to rely on different cognitive control structures that are susceptible to performance breakdown via very different mechanisms (Beilock, Kulp, Holt, & Carr, 2004)—it is important to take a closer look at the processes governing stereotype threat-induced performance decrements across the entire range of skills for which it has been revealed.

Stereotype Threat and Working Memory

In one of the first attempts to directly examine the cognitive mechanisms governing stereotype threat, Schmader and Johns (2003) activated the stereotype that “women are poorer at math than men” in some female participants (stereotype threat condition) or they did not mention gender differences in math to other female participants (control condition). Afterward, women’s working memory capacity and their performance on a difficult math test was measured. Schmader and Johns found that women in the stereotype threat condition showed reduced working memory capacity and poorer math test performance relative to the control group. Furthermore, measured working memory capacity mediated the link between stereotype threat and poorer math performance, supporting a causal relationship. Working memory can be thought of as a short-term memory system that is involved in the control, regulation, and active maintenance of a limited amount of information with immediate relevance to the task at hand (Miyake & Shah, 1999). If working memory capacity is compromised, performance may suffer. Schmader and Johns argued that stereotype threat interferes with performance by reducing the working memory capacity available to attend to a task’s information-processing requirements and to control its execution.

Croizet and colleagues (2004) have come to a similar conclusion regarding the impact of stereotype threat on the cognitive processes required for successful task execution—not by directly measuring working memory but by measuring heart rate variability (HRV), which has been shown to be an indicator of mental workload. Using evidence that a decrease in HRV reflects heightened mental load (Mulder, 1992), Croizet et al. (2004) examined how the performance of undergraduate students for whom a negative intellectual stereotype

existed (in this case, psychology majors with a reputation of lower intelligence in comparison to science majors), and their associated HRV, was affected by a situation that increased the salience of this intelligence-based stereotype. In their study, participants performed the Raven Advanced Progressive Matrices Test (Raven, 1962) under instructions that the test was a “valid measure of general intellectual ability involved in mathematical and logical reasoning” (stereotype threat condition) or a laboratory exercise “not diagnostic of any ability” (no threat condition). When the test was presented as a measure of intellectual ability, students who were psychology majors, and for whom a stereotype about intellectual inferiority was relevant, performed worse than those for whom the same negative stereotype did not apply (i.e., science majors). Moreover, this stereotype-induced performance decrement among psychology majors was mediated by a decrease in HRV. Croizet et al. concluded that increased mental workload under the diagnostic test condition (indicated by a decrease in HRV) was responsible for the poor performance of those individuals susceptible to stereotype threat.

The evidence of increased mental workload among psychology majors seen in Croizet et al.’s (2004) work may reflect, at least in part, the occurrence of negative situation-relevant thoughts that used available cognitive resources over and above those demanded by the problems themselves. This notion is supported by a recent demonstration that negative stereotype-relevant thinking mediates stereotype threat effects in math problem solving. Cadinu, Maass, Rosabianca, and Kiesner (2005) examined women’s math performance under stereotype threat (in which gender differences in math problem solving were made salient) and no threat conditions (in which gender differences in math problem solving were explicitly dispelled). Introducing the negative gender stereotype resulted in a decrease in women’s math performance. Furthermore, this performance decrement was mediated by an increase in domain-specific negative thinking. These findings provide empirical support for Steele, Spencer, and Aronson’s (2002) assertions that stereotype threat is accompanied by “doubts about one’s ability, thoughts about the stereotype” (p. 392). Cadinu et al.’s findings also are consistent with work demonstrating that high pressure-testing situations induce negative thoughts and worries. Such thoughts have been shown to harm performance on tasks such as difficult mathematical problem solving via consumption of working memory resources needed for successful performance (Beilock, Kulp, Holt, & Carr, 2004; Beilock & Carr, 2005).

Taken together, the above work suggests that stereotype threat harms performance by compromising the cognitive control structures, and most likely the working

memory system, needed for the performance of difficult reasoning and math problem-solving tasks. Based on the consistency of this evidence, one might conclude that this is the only mechanism by which stereotype threat can exert its impact. It should be noted, however, that unlike the difficult reasoning and math problem-solving tasks used in the stereotype threat work outlined above, there are other types of skills that may fall prey to stereotype threat effects even though they do not depend heavily on working memory for successful execution. For example, skills that have highly proceduralized components (i.e., skills represented as integrated procedures that run off relatively automatically with minimum intervention from working memory), such as sports skills (e.g., Stone, 2002; Stone et al., 1999), medical routines (e.g., surgery), or musical performances, are thought not to depend on working memory in a manner that should make them susceptible to stereotype threat-induced compromises of this system (Beilock, Carr, MacMahon, & Starkes, 2002). If stereotype threat is revealed for such skills, what would account for its occurrence?

To address this question, it is important to first consider how proceduralized skills succeed and fail. Unlike working memory intensive tasks (e.g., difficult math problem solving), performance on proceduralized tasks has been shown to be harmed by too much attention and not to be harmed (or at times, even benefited) by reductions in attention (e.g., distracting secondary tasks, time pressure). Empirical evidence for this conclusion comes mainly from experiments studying complex sensorimotor skills, such as golf putting, soccer dribbling, and baseball batting, that become proceduralized with extended practice. Such skills are not harmed when working memory capacity is reduced, for example, by a memory load or concurrent task, but they are hurt by directing explicit attention to automatic processes that normally proceed outside working memory's control (Gray, 2004). This is in contrast to working-memory-intensive tasks, such as math problem solving (Beilock, Kulp, et al., 2004).

To illustrate how well-practiced sensorimotor skills are not impaired by distracting situations but are harmed by overattention to aspects of performance that are usually executed in a relatively automatic fashion, Beilock, Carr, et al. (2002) had experienced soccer players dribble a soccer ball while either performing a secondary auditory monitoring task in which they monitored a series of auditory tones for a specific target tone (designed to distract attention away from performance) or a skill-focused task in which individuals monitored the side of their foot that most recently contacted the ball (designed to draw explicit attention to a component process of performance). Auditory monitoring (which

requires substantial cognitive resources) did not harm experienced soccer players' dribbling skill in comparison to a single task (i.e., soccer dribbling alone) baseline condition. However, monitoring the side of the foot that most recently contacted the ball resulted in worse performance in comparison to both the auditory monitoring condition and the single task baseline. Thus for proceduralized skills, performance failure seems to result from overattending to performance (which breaks down actions that are optimally performed in a more automated fashion) instead of from reductions in working memory capacity (because these skills do not rely extensively on these resources). Based on this logic, such skills are either (a) not susceptible to stereotype threat or (b) there are other mechanisms by which stereotype threat may harm their execution.

The Multiple Mechanisms of Stereotype Threat

To uncover additional mechanisms by which skills may fail under threat, we looked to another body of work that explores unwanted skill failure, the "choking under pressure" literature. Here, two theories have been proposed to explain unwanted skill failures.

The first explanation is very much in line with the working memory account of stereotype threat effects outlined above. Distraction theory (Ashcraft & Kirk, 2001; Beilock & Carr, 2005; Beilock, Kulp, et al., 2004; Lewis & Linder, 1997) suggests that pressure-packed situations induce distracting thoughts and worries that compete for the working memory capacity needed for the optimal control of task execution. In essence, negative thoughts and worries about the situation and its outcome utilize the resources that were once solely available to devote to the demanding task at hand. As a result, poor performance ensues.

The second account of skill failure under pressure is quite different but has garnered substantial empirical support. According to explicit monitoring theory (Baumeister, 1984; Beilock & Carr, 2001; Gray, 2004; Lewis & Linder, 1997), performance pressure increases self-consciousness about performing correctly, which in turn induces individuals to increase the attention they devote to controlling step-by-step performance to ensure a positive outcome. Unfortunately, increased attention to proceduralized task control can backfire, disrupting what should have been fluent, automatic execution (Kimble & Perlmutter, 1970; Langer & Imber, 1979; Masters, 1992). Although explicit monitoring is related to other self-focus constructs such as self-awareness (Wicklund & Duval, 1971) or public self-consciousness (Fenigstein, 1979), these latter theories examine attention to the self on a global level, resulting in social

comparison and changes in self-concept, whereas explicit monitoring theory focuses on how attending to step-by-step execution can be counterproductive in real-time task performance. Thus, explicit monitoring theory suggests a very different route by which stereotype threat can impair performance. Although stereotype threat may harm working memory-demanding tasks by reducing the resources available for execution, it may hurt proceduralized skills by prompting too much attention to execution processes that are best run off outside of conscious control.

It might be tempting to conclude from the above descriptions of distraction and explicit monitoring theories that stereotype threat could have one kind of impact on attentional control in working memory intensive skills and another kind of impact on attention in tasks that operate largely outside of conscious awareness. It seems more likely, however, that stereotype threat always exerts at least two different effects—it populates working memory with worries and it entices the performer to try to pay more attention to step-by-step control, resulting in a double whammy. Indeed, recent work has shown that stereotype threat can induce individuals to adopt a prevention focus orientation (Seibt & Förster, 2004), which might not only lead to worry about the possibility of poor performance but also prompt individuals to focus on their own execution to minimize mistakes or adopt avoidant behaviors (Ryan & Ryan, 2005). Thus, stereotype threat may simultaneously affect working memory availability and direct attention in ways that are counterproductive. However, these two effects may be differentially relevant to performance depending on the attentional demands of the task being performed. If a task loads heavily on working memory but does not involve much in the way of proceduralized routines (e.g., difficult and multistep math problem solving), then it will suffer from the reduction of working memory capacity induced by stereotype threat, but it will be relatively indifferent to the attempt to focus what attention remains more specifically on step-by-step control that also is induced by stereotype threat. Conversely, if a task relies heavily on proceduralized routines but puts little stress on working memory (e.g., well-learned sensorimotor skills), then that task will suffer from stereotype threat because of the stereotype threat-induced shift of attention to step-by-step control and not because the overall capacity of working memory has been reduced.

As outlined above, there is evidence in support of the notion that stereotype threat harms performance in working memory intensive tasks such as math problem solving by reducing or consuming the working memory resources required for performance (e.g., Schmader &

Johns, 2003). To determine whether stereotype threat might operate via multiple mechanisms of failure, the current work explored the causal mechanisms of stereotype threat in a very different type of skill—a skill that should be harmed by explicit attention to execution rather than by reductions in the working memory resources available for performance.

We chose the sensorimotor task of golf putting as our test bed, partly because it is a complex task placing great demands on sensorimotor control and partly because it has already been used in stereotype threat studies (e.g., Stone, 2002; Stone et al., 1999). Furthermore, the cognitive mechanisms involved in golf putting have been extensively studied and are reasonably well understood (e.g., Beilock, Bertenthal, et al., 2004; Beilock, Carr, et al., 2002; Lewis & Linder, 1997; Perkins-Ceccato, Passmore, & Lee 2003). Golf putting also reveals underlying processes similar to other proceduralized skills (e.g., Beilock, Carr, et al., 2002; Gray, 2004). All of this makes studying golf putting as a target skill especially beneficial. To the extent that proceduralized skills such as high-level putting fall prey to stereotype threat effects even though they do not rely heavily on working memory resources (Beilock, Carr, et al., 2002; Gray, 2004), this would demonstrate that there are multiple mechanisms through which stereotype threat can exert its impact.

Moreover, if support for an explicit monitoring explanation of stereotype threat is found, it would provide a bridge between the stereotype threat and choking under pressure literatures. That is, demonstrating similar mechanisms of failure in proceduralized skills—whether such failures are induced by situational pressure to perform at a high level or by fear of confirming a negative performance stereotype about one's social group—would highlight the similarities in the processes governing these two types of skill failure that surprisingly have been studied largely in isolation (Beilock & McConnell, 2004).

Current Experiments

Experiment 1 sought to demonstrate stereotype threat specifically among extremely accomplished golfers. We chose expert golfers to be our participants to ensure that their golf skills would be relatively proceduralized (via the countless hours of practice that they have put into honing their skill). Furthermore, past research has demonstrated that stereotype threat effects in sport are moderated by how important performance is to athletes' sense of self-worth (i.e., greater stereotype threat revealed for those who perceive sports as more important; Stone, 2002; Stone et al., 1999).

Thus, it seems likely that high-level athletes, who place considerable importance on success, would be especially likely to increase the attention they devote to execution to ensure a positive outcome in response to stereotype threat (Baker & Horton, 2003; Beilock & McConnell, 2004). In sport, then, those with the highest skill level and the most invested in personal success may be the most susceptible to failure. Although the impact of skill level and investment would be similar to stereotype threat findings in academic tasks (Steele, 1997), the mechanism of failure might be quite different. Experiments 2 and 3 explicitly tested this differential mechanism more directly.

Experiment 2 employed the logic that if stereotype threat in proceduralized skills such as golf putting results from explicit attention to performance rather than by consumption of working memory resources, then adding a working memory demanding secondary task to golf putting performance under stereotype threat (e.g., an auditory monitoring task in which individuals are required to constantly monitor a series of words for a specified target word while performing the putting task) should actually improve performance relative to stereotype threat alone. This manipulation was first employed by Lewis and Linder (1997) in an attempt to demonstrate that pressure-induced performance decrements in a well-learned golf putting skill could be alleviated by the introduction of a secondary task (in this case, a backward counting task) that served to draw attention away from performance execution. Thus, adding a dual task to stereotype threat may actually help performance because the secondary task would prevent stereotype-induced attention from being devoted to experts' automated or proceduralized performance processes. Such a result would be completely at odds with a working memory account of stereotype threat, which would suggest that adding a secondary task that requires working memory under stereotype threat should hurt performance even more than stereotype threat alone.

It is possible, however, that giving individuals a secondary task to perform while under stereotype threat might serve not to draw attention away from putting performance that is best run off outside of conscious awareness but rather it might draw attention away from the stereotype itself. To ensure that this was not the case, in Experiment 3, we changed the content of the dual task following the introduction of stereotype threat to ensure that the dual task served to keep the stereotype active in memory while golfers performed the putting task. Moreover, we changed the stereotype threat manipulation in this experiment to extend the generalizability of the current work.

EXPERIMENT 1

Male athletes who were expert golfers performed a series of golf putts on an indoor putting green both before and after receiving either a negative stereotype about their golf putting performance or control information. In this work, we used men because our stereotype threat manipulation involved gender and because the vast majority of expert golfers in our sample population were men. A working memory account of failure under stereotype threat would propose that stereotype threat operates by reducing the working memory available for execution. Thus, expert performance, thought to run largely outside of working memory (Anderson, 1993; Beilock, Carr, et al., 2002; Fitts & Posner, 1967; Gray, 2004), should not be harmed by stereotype threat. In contrast, if stereotype threat harms execution by inducing attention and explicit monitoring of performance, then the proceduralized performance of experts should show stereotype-induced decrements.

Method

Participants. Forty men participated and were randomly assigned to either a control group ($n = 22$) or a stereotype threat group ($n = 18$). To participate, all individuals were required to have 2 or more years of high school/college varsity golf experience or a Professional Golfers' Association (PGA) handicap ≤ 8 . In addition, to help ensure that the golfers held strong perceptions of their athletic ability and the importance of athletics, participants also were required to have 2 or more years of non-golf-related high school varsity sport experience or equivalent. Finally, we employed a cut-off for study participation for individuals who reported poor athletic ability or low importance of this ability. Individuals responded to a 9-point question assessing the importance of athletic ability and a 10-point question assessing athletic ability relative to other college students, based on Pelham and Swann's (1989) Self-Attributes Questionnaire, which assesses perceptions of ability in different domains (e.g., social competence, athletic ability). A mean score of 5 or higher on these two questions was required for participation.

Procedure. Participants were run individually by a male or female experimenter.¹ After providing informed consent and completing a questionnaire about previous sports experience, participants were introduced to the putting task. Based on the procedures of Beilock and Carr (2001), all participants performed the same putting task on a carpeted indoor green (3 m \times 3.7 m) using a standard golf putter and ball. Individuals putted a golf ball as accurately as possible to a target on which

the ball was supposed to stop. Individuals putted in a random order from nine different starting locations, with three locations at three different distances (120 cm, 140 cm, 150 cm) from the target. Before each putt, a light indicated the location from which participants were to take their next putt. After each putt, and while the participant set up for the next putt, the experimenter measured putting accuracy by the distance in cm away from the center of the target that the ball had stopped.

Participants first performed 18 practice putts to familiarize themselves with the putting task. Next, they completed an 18 putt pretest measure of putting performance (pretest). Immediately following the pretest, participants read more about the purpose of the experiment. All participants read that the study involved researching golf putting ability. Participants in the control condition also read that this research was investigating individual differences in golf putting performance. In contrast, participants in the stereotype threat group read that in our laboratory, we have been researching individual differences in golf putting ability and, in particular, focusing on gender. Specifically, they were told that women actually tend to perform better than men on our putting task. Stereotype threat participants also read that this gender difference in putting was further supported by a recent examination of PGA and Ladies Professional Golf Association (LPGA) playing statistics demonstrating that women tend to be better short game players than men.

Next, all participants performed two more blocks of 18 putts, which were combined to form the posttest performance measure. Individuals were then fully debriefed.

Results

The mean distance (cm) from the target that the ball stopped after each putt in the pretest and posttest served as the measure of that condition's putting performance. Putts greater than 3 standard deviations above each individual's pretest and posttest averages were considered outliers and removed, resulting in the elimination of 15 individual putt outliers (< 1% of all putts).

We began by comparing putting performance preceding our manipulation (i.e., the pretest) to putting performance following the manipulation as a function of whether individuals received the negative stereotype about their golf putting ability. A 2 (group: control, stereotype threat) \times 2 (test: pretest, posttest) ANOVA revealed a significant Group \times Test interaction, $F(1, 38) = 5.39, p < .03, \eta_p^2 = 0.12$.

As seen in Figure 1, there was no difference in putting performance as a function of group in the pretest, $F < 1$. However, in the posttest, the stereotype threat group

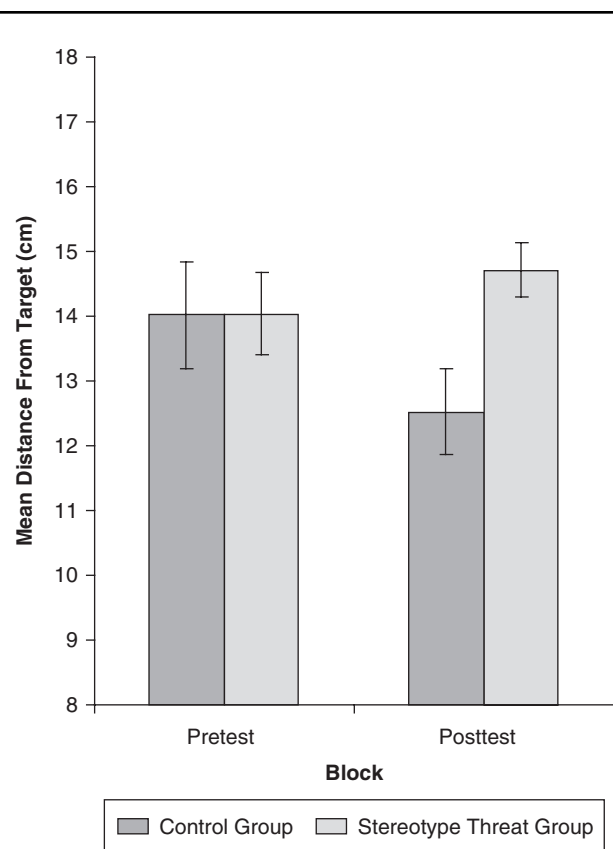


Figure 1 Mean distance from the target (cm) that the ball stopped after each putt in the pretest and posttest for expert golfers under stereotype threat and control conditions in Experiment 1.

NOTE: Error bars represent standard errors.

performed significantly worse than the control group, $F(1, 38) = 7.01, p < .02, d = 0.87$.

Furthermore, golfers in the control group significantly improved in putting accuracy from the pretest to the posttest, $t(21) = 2.24, p < .04, d = 0.42$. In contrast, expert golfers in the stereotype threat group got worse, although this decrement was not significant, $t(17) = 1.1, ns$.

Discussion

Golfers in the control and stereotype threat groups did not differ in putting accuracy in the pretest, but following the stereotype threat manipulation, stereotype threat group participants performed significantly worse than their control group counterparts. According to a working memory account of stereotype threat, the introduction of a negative stereotype harms execution by reducing the working memory capacity available for performance. Using this logic, expert golf putting performance, demonstrated to run largely outside of working

memory, should not be harmed (Beilock & Carr, 2001). According to explicit monitoring theory, however, stereotype threat may induce experts to increase the amount of attention and working memory capacity they devote to controlling performance. This increased attention should disrupt fluent, proceduralized performance (Beilock, Carr, et al., 2002). The current pattern of results supports the explicit monitoring account.

It should be noted, however, that expert putting performance revealed stereotype threat effects in terms of a lack of improvement in putting accuracy from the pretest to the posttest rather than a significant decrease in performance relative to the pretest. Because our putting task required golfers to land a ball on a target rather than in a hole, experts were most likely adjusting to the novel demands of our putting task and thus continuing to improve in putting accuracy throughout the course of the experiment. Comments from experts performing this task in previous studies support such a contention. For example, many golfers have remarked that it was a bit odd at first to “stop the ball on the tape and not go through it” but that if one just thought of the green “as very fast,” they could perform successfully. Consequently, rather than stereotype threat significantly harming performance relative to a stable and unchanging baseline state, it prevented improvement in performance relative to a non-stereotype-threat control group. Indeed, the performance of experts under stereotype threat in the posttest was significantly worse than the performance of control group experts, revealing the detrimental consequences of negative group stereotypes.

Still, the fact that experts needed to adapt to the novel aspects of our task might cause one to worry that our experts were exerting some step-by-step control over their performance in a working memory intensive fashion. Were this the case, then reduction of working memory capacity by worries about the situation might remain a viable explanation of the results reported so far. The feasibility of a working memory explanation was tested further in the next experiment.

If stereotype threat harms performance by prompting explicit attention to execution in proceduralized sensorimotor skills such as golf putting, then it follows that drawing attention away from performance should reduce this type of maladaptive attentional control. That is, the introduction of a concurrent task should lessen stereotype-induced overattention to performance processes that are best left outside conscious awareness, thereby reducing the detrimental impact of a negative performance stereotype. However, if stereotype threat is acting by reducing working memory capacity, then adding an even greater burden on working memory capacity via secondary task demands

should harm performance even more. Experiment 2 tested these contrasting predictions.

EXPERIMENT 2

Male athletes who were expert golfers performed a series of golf putts under both single-task and dual-task conditions, both before and after receiving the same negative stereotype as Experiment 1. As mentioned above, high-level sensorimotor skills are thought to be governed by procedural knowledge that does not require on-line control and operates largely outside of working memory (Beilock & Carr, 2001). Consequently, attention can be devoted to a secondary task without significantly disrupting primary task performance (Allport, Antonis, & Reynolds, 1972; Beilock, Carr, et al., 2002; Ford, Hodges, & Williams, 2005; Gray, 2004). Thus, under no threat conditions, expert golf putting should not be significantly harmed by the addition of a concurrent task. More interesting, however, under stereotype threat, the dual-task condition may actually improve performance if the secondary task prevents stereotype-induced attention from being devoted to experts' automated or proceduralized performance processes. Such a result would be completely at odds with a working memory account of stereotype threat because this account suggests that adding a secondary task that occupies working memory under stereotype threat should hurt performance even more than stereotype threat alone.

Method

Participants. Participants ($N = 23$) were male athletes who are expert golfers, using the same criteria as Experiment 1.²

Procedure. Participants performed a similar putting task on the same green used in Experiment 1. All participants again completed one block of 18 practice putts to familiarize them with the putting green. These practice putts were followed by an initial 18 putts in a single-task condition and an initial 18 putts in a dual-task condition (described below), with order counterbalanced across participants. All participants then read the same stereotype threat manipulation used in Experiment 1. Thus, in Experiment 2, stereotype threat was manipulated within subjects.

Following the introduction of stereotype threat, participants completed a second block of 18 putts under the single-task condition and a second block of 18 putts under the dual-task condition. Again, the order of these putting blocks was counterbalanced across participants. To reinforce the salience of the negative performance stereotype prior to the second block of putts

following the stereotype threat manipulation, individuals were given a short questionnaire to complete in between the first and second block of putts following the stereotype threat manipulation. To the participant, the questionnaire appeared to be a short memory test regarding the purpose of the study. Participants responded to seven questions, four of which concerned the stereotype threat manipulation (e.g., Previous research “has demonstrated that ___ golfers often show superior short game play to ___ golfers. Please fill in the blank.” The answer here is “female, male”). All participants responded correctly to all of the questions regarding the stereotype threat manipulation. Following the putting task, participants were fully debriefed.

Single-task condition. In the single-task condition, participants performed the putting task in a normal, single-task environment (as in Experiment 1).

Dual-task condition. The dual-task condition involved putting while simultaneously listening to a series of spoken single-syllable nouns being played from a tape recorder. Participants were instructed to monitor the words carefully, and each time they heard a specified target word (either *thorn* or *dean*), to repeat the word aloud. For each participant, one of the target words was used in the first dual-task putting block (before the stereotype threat manipulation) and the other target word was used in the second dual-task putting block (after the stereotype threat manipulation). The order of the target words was counterbalanced across participants.

Words were presented at a randomly determined moment once within 2-s time intervals. The target word was presented randomly, approximately once in every four words. The random placement of the words within the 2-s time intervals, as well as the random embedding of the target word among the filler words, was designed to prevent the golfers from anticipating target word presentation.

Results

Putting performance. As in Experiment 1, putting accuracy was measured by the distance (in cm) away from the center of the target that the ball stopped after each putt. The mean distance from the target of the 18 putts in each block (i.e., the single-task and dual-task blocks before stereotype threat and the single-task and dual-task blocks following stereotype threat) served as the measure of that condition's putting performance. As in Experiment 1, putts greater than 3 standard deviations above each individual's mean for each block were removed (5 putts were removed in total, < 1% of all putts).

A 2 (condition: single task, dual task) \times 2 (stereotype threat: before, after) within-subjects ANOVA on putting accuracy revealed a significant Stereotype Threat \times

Putting Condition interaction, $F(1, 22) = 6.03$, $p < .03$, $\eta_p^2 = 0.22$, which is shown in Figure 2. Prior to the introduction of the negative performance stereotype, there was no significant difference between single-task and dual-task putting, $t(22) = 0.99$, *ns*. This result replicates previous work with sensorimotor skills demonstrating that high-level performances are not significantly compromised by situations that draw attention away from execution (Allport et al., 1972; Beilock, Carr, et al., 2002; Beilock, Wierenga, & Carr, 2002; Ford et al., 2005; Gray, 2004). What small (but nonsignificant) difference there was indicated a bit of a decline in performance under dual-task conditions, consistent with the notion that stopping the ball on the target required some adaptation, but not enough to overcome the proceduralized skills of these expert golfers. In contrast, following the introduction of stereotype threat, participants performed significantly worse in the single-task than in the dual-task condition, $t(22) = 2.32$, $p < .03$, $d = 0.52$. Furthermore, replicating Experiment 1, the introduction of the negative stereotype ameliorated any improvement in performance from the single-task condition before the introduction of stereotype threat to the single-task condition following stereotype threat. Based on the putting performance of the control group in Experiment 1 (i.e., those not under stereotype threat), single-task putting performance should be expected to improve somewhat as golfers gain more experience with the particular features of the putting task being performed in the current work. Stereotype threat, however, retarded this improvement.

Word-monitoring performance. Each instance in which individuals failed to identify a target word was recorded. In total, only 6 target words failed to be identified across all of the presentations given to the 23 participants in both dual-task conditions. On average, participants heard 18.6 words (4.7 of which were target words and the rest filler words in each dual-task condition). Thus, the failure rate for target words was a very low 2.8%, indicating that participants were extremely accurate at identifying target words in the dual-task conditions and thus very attentive to the secondary task. This suggests that the pattern of putting performance data reported above is not the result of a trade-off with word-monitoring task performance.

Discussion

Expert golfers performed a series of putts under single-task and dual-task conditions both before and after receiving negative stereotype information. Prior to the introduction of the negative stereotype, putting performance between the single-task and dual-task conditions did not significantly differ. Because experts do not have

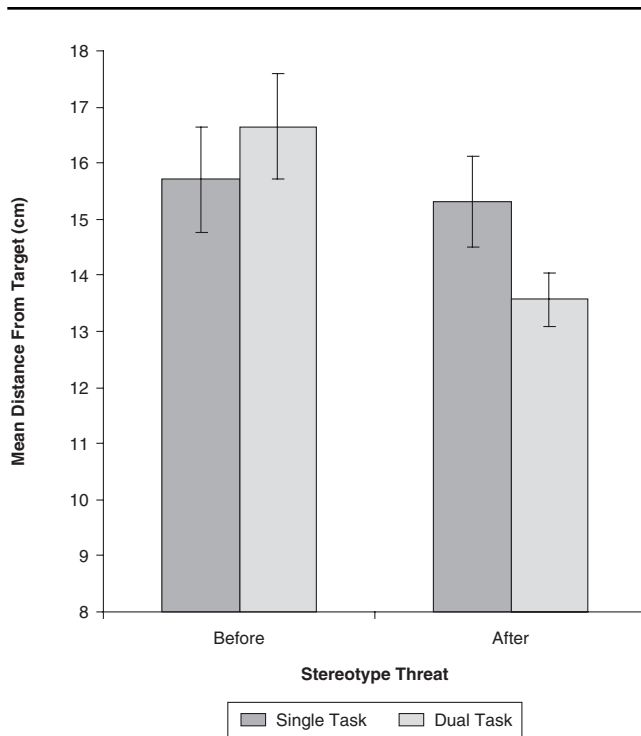


Figure 2 Mean distance from the target (cm) that the ball stopped after each putt in the single-task and dual-task conditions both before and after the introduction of stereotype threat in Experiment 2.

NOTE: Error bars represent standard errors.

to attend to every step of performance, skill execution does not suffer much when attention is drawn away from performance by secondary task demands. In contrast, following the introduction of the negative stereotype, experts' performance was significantly more accurate in the dual-task than single-task condition. If stereotype threat impairs performance by prompting explicit attention to execution in well-learned proceduralized skills such as golf putting, then drawing attention away from performance (e.g., through the introduction of a secondary task) should reduce this type of maladaptive attentional control. This is precisely the pattern of data that was observed.

The results of Experiments 1 and 2 diverge from previous stereotype threat work, which suggests that negative stereotypes harm performance in complex cognitive tasks such as math problem solving by reducing the resources necessary for performance (Cadinu et al., 2005; Schmader & Johns, 2003). Instead, the current work suggests that in high-level proceduralized skills such as golf putting, stereotype threat-induced performance decrements result from too much attention to execution rather than too little. Indeed, adding a secondary task under stereotype threat improved performance

rather than hurt it. We propose that these different mechanisms of failure are due to differences in the representation of working-memory-intensive and proceduralized skills. Although skills such as difficult math problem solving appear to involve heavy on-line processing demands and thus are harmed when such demands are not met, well-learned skills that become proceduralized with practice run largely outside of working memory and thus are harmed by attempts to bring performance back under step-by-step control (for further discussion, see Beilock, Kulp, et al., 2004).

It is possible, though, that the current results deviate from previous work not because of the type of task under investigation but because of the type of stereotype that was used. Unlike previous work that has used culturally held stereotypes to induce stereotype threat, Experiments 1 and 2 provided participants with a fictitious negative stereotype about performance. However, whether stereotypes are held by others or whether they are veridical is not as important as whether individuals perceive them to be diagnostic of group members' characteristics (Hamilton & Sherman, 1994). All that appears necessary for stereotype threat to emerge is that an individual is both aware of the stereotype and aware that the task they are performing is diagnostic of the ability in question (Kray, Galinsky, & Thompson, 2002). To the extent that stereotype threat effects are realized in Experiments 1 and 2, this should be, if anything, more difficult to obtain with fictitious rather than with culturally held or normatively shared stereotypes. Nonetheless, to ensure that our results were not due to this difference, Experiment 3 replicated Experiment 2 using a culturally shared stereotype concerning the relation between race and athletic ability, borrowed from Stone et al. (1999).

Another possible reason that our expert golfers, following stereotype threat, putted better under dual-task than single-task conditions was not because the dual task diverted attention away from execution but because the dual task diverted attention away from the stereotype. If the latter were true, we would not have evidence that the secondary task prevented stereotype-induced attention to execution (supporting an explicit monitoring account of stereotype threat). All we would have shown is that the secondary task prevented golfers from thinking about the stereotype.

Indeed, there is evidence that diverting attention from the stereotype may well serve to reduce stereotype threat effects. Stone et al. (1999) found that providing golfers with a misattribution cue for their arousal under stereotype threat (i.e., an uncomfortable lab performance space) resulted in better performance than when such a cue was not provided. These researchers suggested that the misattribution cue drew performers' attention to the

lab space, thus preventing them from thinking about the negative stereotype associated with their group identity (see also Ben-Zeev, Fein, & Inzlicht, 2005).

Thus, in Experiment 3, to ensure that the secondary task did not merely prevent golfers from thinking about the stereotype, we changed the content of the dual task following the introduction of stereotype threat. Specifically, the dual task consisted of race-related words designed to keep the stereotype active in memory while golfers performed the putting task. Approximately half of the words (including the target word) were race-related. Previous research has demonstrated that expert golfers attend to words heard in dual-task putting and word-monitoring situations enough to be able to detect target words and to recognize distractor words on later memory tests (Beilock, Wierenga, & Carr, 2002). Thus, presenting racially related words as targets and distractors should serve to maintain stereotype activation while experts perform the putting task. Indeed, if experts are performing the secondary task at all (and given the low error rates, there is ample evidence they are), the mere identification of the race-related target word should serve to keep the stereotype active in memory.

EXPERIMENT 3

Method

Participants. Participants ($N = 17$) were Caucasian male athletes who were expert golfers using the same criteria as the first two experiments.

Procedure. The design was exactly the same as Experiment 2, with the following exceptions: First, participants putted from five starting locations rather than nine and there were 20 putts in each block rather than 18. These two changes occurred because a different putting green was used. Also, the stereotype given to participants was a culturally shared stereotype (adapted from Stone et al., 1999) designed to make salient the stereotype that Whites do not have as much natural athletic ability as Blacks. Specifically, in the stereotype threat manipulation, individuals were told that natural athletic ability is the best predictor of performance success in golf, and as a result, certain minority groups may be especially predisposed to excel in golf (including our putting task). Finally, to ensure that our secondary task did not distract attention away from the stereotype, half of the words that participants monitored in the dual-task condition following stereotype threat (but not before) were race-related. Specifically, the target word was *White*, which occurred on 25% of the presentations. Also, to ensure that *White* was viewed in a racial context (see Wittenbrink, Judd, & Park, 1997), 25% of the

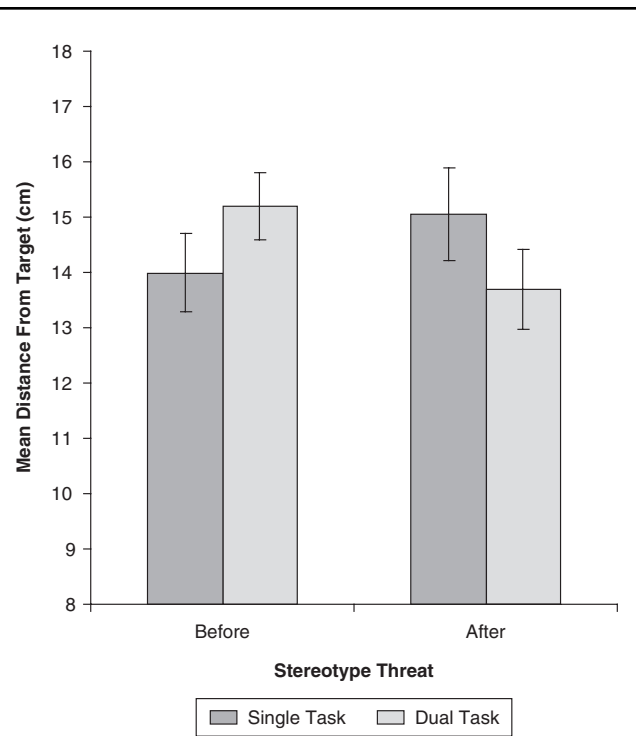


Figure 3 Mean distance from the target (cm) that the ball stopped after each putt in the single-task and dual-task conditions both before and after the introduction of stereotype threat in Experiment 3.

NOTE: Error bars represent standard errors.

nontarget words presented were associated with either Black or White stereotypes (e.g., jazz, rich), as established by pretesting.

Results

Putting performance. As in Experiments 1 and 2, putts greater than 3 standard deviations above each individual's mean for each block were removed. Six total putts were removed (< 1% of the putts).

A 2 (putting condition: single task, dual task) \times 2 (stereotype threat: before, after) within-subjects ANOVA on putting accuracy revealed a significant Stereotype Threat \times Condition interaction, $F(1, 16) = 5.49$, $p < .04$, $\eta_p^2 = 0.26$. As can be seen in Figure 3, and replicating Experiment 2, prior to the introduction of the negative performance stereotype, there was no significant difference between single-task and dual-task performance, $t(16) = 1.47$, *ns*, and the difference in absolute accuracies that did occur was in the direction of slightly worse performance in the dual-task condition. In contrast, following the introduction of stereotype threat, golfers performed significantly worse in the single-task than dual-task condition, $t(16) = 2.31$, $p < .04$, $d = 0.41$.

Word-monitoring performance. As in Experiment 2, each instance in which individuals failed to identify a target word was recorded. The failure rate for target word identification was a very low 3.6% across both dual-task conditions. Again, individuals appeared to be attentive to the secondary task, and thus, putting performance does not seem to be accountable for by a trade-off with the secondary task.

Discussion

Prior to encountering the negative stereotype, expert performance across the single-task and dual-task conditions did not significantly differ. Yet, following the presentation of the negative performance stereotype, performance was significantly more accurate in the dual-task than single-task condition. These results replicate Experiment 2 using a culturally held stereotype and are not likely to be due to the secondary task diverting experts' attention away from the stereotype, given that the secondary task itself required processing and responding to stereotype-relevant information.

GENERAL DISCUSSION

Previous work has shown that stereotype threat harms working-memory-intensive tasks such as math problem solving by reducing the working memory capacity available for performance (Schmader & Johns, 2003). The current work demonstrates that this is not the only mechanism by which stereotype threat exerts its impact. Rather, in well-learned skills that do not rely heavily on working memory for successful performance (e.g., proceduralized sensorimotor skills), stereotype threat harms execution by inducing too much attention to execution rather than too little. These findings are consistent with work on "choking under pressure" (Beilock & Carr, 2001; Gray, 2004), suggesting commonalities between stereotype threat and performance pressure as factors influencing cognitive processes.

In addition, the current work suggests that there are multiple mechanisms by which unwanted performance decrements can occur, and how task performance is represented and implemented will determine the way in which a skill fails. Because tasks such as difficult math problems rely heavily on working memory for successful execution, they suffer from the reductions in available working memory capacity that stereotype threat induces (Croizet et al., 2004; Schmader & Johns, 2003). However, for tasks that rely heavily on proceduralized routines and are not heavily working memory dependent, such as the putting task used in the current work, reductions in working memory do not harm execution. Rather, stereotype-threat-induced shifts of attention to

step-by-step control disrupts execution processes best left outside of working memory. Given the fact that stereotype threat likely both consumes working memory and prompts individuals to try and control performance, tasks that concurrently load on working memory and rely on proceduralized skills might be susceptible to both effects at once.

It should be noted that in the current work prior to stereotype threat introduction in Experiments 2 and 3, there were modest (although not significant) differences between single and dual-task performance. This is likely due to our somewhat novel putting task that involved landing a ball on a target rather than in a hole. Adjusting to these new task constraints probably required some cognitive resources, and thus, performance was slightly reduced under dual-task conditions. This small dual-task decrement prior to stereotype threat, however, was turned into a significant dual-task advantage following stereotype threat. Such results not only provide strong support for explicit monitoring in stereotype threat but also demonstrate the impressive impact of occupying attention with a secondary task after exposing individuals to stereotype threat—a finding with important implications for counteracting stereotype threat effects. Specifically, for skills that have highly proceduralized components such as sports skills, medical routines (e.g., surgery), or musical performances, diverting attention to a secondary attention-demanding task (e.g., counting backward) in threatening situations may prevent stereotype-induced attention to skill processes and procedures that are best executed without conscious control.

One might wonder whether the current work demonstrates expectancy effects (e.g., Rosenthal & Jacobson, 1966) rather than stereotype threat. That is, maybe the golfers in the current experiments exerted less effort in their performance because the negative stereotype they received suggested that the experimenter was biased against them. Although all stereotypes involve expectancies, such an explanation for the above findings would have difficulty accounting for those situations where the introduction of negative performance stereotypes did not harm performance. For example, in Experiments 2 and 3, the dual-task condition alleviated stereotype-induced performance decrements. If participants merely exerted less effort because they perceived that the experimenter was biased against them, then there is no reason why performance under stereotype threat should be better in the dual-task in comparison to the single-task condition. Rather, the findings from the current three experiments seem most consistent with the idea that stereotype threat harms performance in proceduralized skills by prompting attention to execution. This type of maladaptive explicit

attention can be eliminated by the addition of secondary task demands.

One also might question our choice of secondary tasks and ask whether a secondary task that was largely motorically based (rather than verbal) would produce similar effects. The purpose of the secondary task was to utilize attentional resources that might otherwise, under stereotype threat, be allocated to controlling step-by-step execution in a maladaptive manner. Whether one performs a verbal secondary auditory monitoring task that consumes attentional resources or a more motoric task such as tapping their big toe to a metronome in a manner that also is attention demanding should not matter, as long as the task prevents the explicit attentional control of proceduralized performance processes while at the same time allowing the performer to execute the physical movements necessary to complete the task.

Nonetheless, are there other accounts of stereotype activation that could anticipate the above findings? For example, can ideomotor theories (Wheeler & Petty, 2001) explain our results? According to an ideomotor explanation, the golfers in the current work may have performed more poorly under stereotype threat than control conditions because the negative stereotype presented to individuals activated social categories associated with poor performance that, in turn, automatically activated demonstrations of impaired performance (see Bargh & Chartrand, 1999). Although such a direct priming account might predict stereotype threat effects under single-task conditions, it cannot readily account for why golfers performed significantly better in a dual-task stereotype threat condition in comparison to stereotype threat alone. That is, if stereotype threat harms well-learned putting by directly and automatically activating concepts of poor performance, this should not be diminished when a secondary task is added to performance under stereotype threat, especially when the secondary task is designed to keep the poor performance stereotype active in memory.

To conclude, the current work underscores the importance of understanding the mechanisms involved in stereotype threat for the purpose of alleviating such unwanted skill failures. For example, if one considered remedies for stereotype threat based solely on the presumption that stereotype threat operates by consuming processing resources, one might suggest that stereotype threat can be reduced by increasing attention and working memory devoted to performance. Although such a prescription might reduce stereotype threat decrements for members of stigmatized groups on difficult college entrance exams, these recommendations would most likely backfire for stigmatized group members performing more proceduralized skills in an

orchestra pit, an operating theater, or on an athletic field. Thus, in some activities, distraction has its benefits.

NOTES

1. The sex of the experimenter was counterbalanced across conditions and revealed no significant effects in any of the experiments.
2. Two participants were dropped from Experiment 2 because they failed to identify at least 50% of the targets in either dual-task blocks. No participants failed this criterion in Experiment 3.

REFERENCES

- Allport, D. A., Antonis, B., & Reynolds, P. (1972). On the division of attention: A disproof of the single channel hypothesis. *Quarterly Journal of Experimental Psychology*, *24*, 225-235.
- Anderson, J. R. (1993). *Rules of mind*. Hillsdale, NJ: Lawrence Erlbaum.
- Ashcraft, M. H., & Kirk, E. P. (2001). The relationships among working memory, math anxiety, and performance. *Journal of Experimental Psychology: General*, *130*, 224-237.
- Baker, J., & Horton, S. (2003). East African running dominance revisited: A role for stereotype threat? *British Journal of Sports Medicine*, *37*, 553-555.
- Bargh, J. A., & Chartrand, T. L. (1999). The unbearable automaticity of being. *American Psychologist*, *54*, 462-479.
- Baumeister, R. F. (1984). Choking under pressure: Self-consciousness and paradoxical effects of incentives on skillful performance. *Journal of Personality and Social Psychology*, *46*, 610-620.
- Beilock, S. L., Bertenthal, B. L., McCoy, A. M., & Carr, T. H. (2004). Haste does not always make waste: Expertise, direction of attention, and speed versus accuracy in performing sensorimotor skills. *Psychonomic Bulletin & Review*, *11*, 373-379.
- Beilock, S. L., & Carr, T. H. (2001). On the fragility of skilled performance: What governs choking under pressure? *Journal of Experimental Psychology: General*, *130*, 701-725.
- Beilock, S. L., & Carr, T. H. (2005). When high-powered people fail: Working memory and "choking under pressure" in math. *Psychological Science*, *16*, 101-105.
- Beilock, S. L., Carr, T. H., MacMahon, C., & Starkes, J. L. (2002). When paying attention becomes counterproductive: Impact of divided versus skill-focused attention on novice and experienced performance of sensorimotor skills. *Journal of Experimental Psychology: Applied*, *8*, 6-16.
- Beilock, S. L., Kulp, C. A., Holt, L. E., & Carr, T. H. (2004). More on the fragility of performance: Choking under pressure in mathematical problem solving. *Journal of Experimental Psychology: General*, *133*, 584-600.
- Beilock, S. L., & McConnell, A. R. (2004). Stereotype threat and sport: Can athletic performance be threatened? *Journal of Sport and Exercise Psychology*, *26*, 597-609.
- Beilock, S. L., Wierenga, S. A., & Carr, T. H. (2002). Expertise, attention, and memory in sensorimotor skill execution: Impact of novel task constraints on dual-task performance and episodic memory. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, *55*, 1211-1240.
- Ben-Zeev, T., Fein, S., & Inzlicht, M. (2005). Arousal and stereotype threat. *Journal of Experimental Social Psychology*, *41*, 174-181.
- 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.
- Croizet, J., Despres, G., Gauzins, M., Huguet, P., Leyens, J., & Meot, A. (2004). Stereotype threat undermines intellectual performance by triggering a disruptive mental load. *Personality and Social Psychology Bulletin*, *30*, 721-731.
- Fenigstein, A. (1979). Self-consciousness, self-attention, and social interaction. *Journal of Personality and Social Psychology*, *37*, 75-86.
- Fitts, P. M., & Posner, M. I. (1967). *Human performance*. Belmont, CA: Brooks/Cole.
- Ford, P., Hodges, N. J., & Williams, A. M. (2005). Online attentional-focus manipulations in a soccer-dribbling task: Implications for

- the proceduralization of motor skills. *Journal of Motor Behavior*, 37, 386-394.
- Gray, R. (2004). Attending to the execution of a complex sensorimotor skill: Expertise differences, choking and slumps. *Journal of Experimental Psychology: Applied*, 10, 42-54.
- Hamilton, D. L., & Sherman, J. W. (1994). Stereotypes. In T. K. Srull & R. S. Wyer (Eds.), *Handbook of social cognition* (2nd ed., Vol. 2, pp. 1-68). Hillsdale, NJ: Lawrence Erlbaum.
- Johns, F., 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.
- Kimble, G. A., & Perlmutter, L. C. (1970). The problem of volition. *Psychological Review*, 77, 361-384.
- Kray, L. J., Galinsky, A. D., & Thompson, L. (2002). Reversing the gender gap in negotiations: An exploration of stereotype regeneration. *Organizational Behavior and Human Decision Processes*, 87, 386-409.
- Langer, E., & Imber, G. (1979). When practice makes imperfect: Debilitating effects of overlearning. *Journal of Personality and Social Psychology*, 37, 2014-2024.
- Lewis, B., & Linder, D. (1997). Thinking about choking? Attentional processes and paradoxical performance. *Personality and Social Psychology Bulletin*, 23, 937-944.
- Masters, R. S. (1992). Knowledge, knerves and know-how: The role of explicit versus implicit knowledge in the breakdown of a complex motor skill under pressure. *British Journal of Psychology*, 83, 343-358.
- Miyake, A., & Shah, P. (1999). Toward unified theories of working memory: Emerging general consensus, unresolved theoretical issues, and future research directions. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 442-481). New York: Cambridge University Press.
- Mulder, L.J.M. (1992). Measurement and analysis methods of heart rate and respiration for use in applied environments. *Biological Psychology*, 34, 205-236.
- Pelham, B. W., & Swann, W. B., Jr. (1989). From self-conceptions to self-worth: On the sources and structure of global self-esteem. *Journal of Personality and Social Psychology*, 57, 672-680.
- Perkins-Ceccato, N., Passmore, S. R., & Lee, T. D. (2003). Effects of focus of attention depend on golfers' skill. *Journal of Sports Sciences*, 21, 593-600.
- Raven, J. C. (1962). *Advanced progressive matrices: Sets I and II*. London: H. K. Lewis.
- Rosenthal, R., & Jacobson, L. (1966). Teachers' expectancies: Determinants of pupils' IQ gains. *Psychological Reports*, 19, 115-118.
- Ryan, K. E., & Ryan, A. M. (2005). Psychological processes underlying stereotype threat and standardized math test performance. *Educational Psychologist*, 40, 53-63.
- Schmader, T., & Johns, M. (2003). Converging evidence that stereotype threat reduces working memory capacity. *Journal of Personality and Social Psychology*, 85, 440-452.
- Seibt, B., & Förster, J. (2004). Stereotype threat and performance: How self-stereotypes influence processing by inducing regulatory foci. *Journal of Personality and Social Psychology*, 87, 38-56.
- Spencer, S. J., Steele, C. M., & Quinn, D. M. (1999). Stereotype threat and women's math performance. *Journal of Experimental Social Psychology*, 35, 4-28.
- Steele, C. M. (1997). A threat in the air: How stereotypes shape intellectual identity 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. P. Zanna (Ed.), *Advances in experimental social psychology* (pp. 379-440). Amsterdam: Academic Press.
- 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.
- Wheeler, S. C., & Petty, R. E. (2001). The effects of stereotype activation on behavior: A review of possible mechanisms. *Psychological Bulletin*, 127, 797-826.
- Wicklund, R. A., & Duval, S. (1971). Opinion change and performance facilitation as a result of objective self-awareness. *Journal of Experimental Social Psychology*, 7, 319-342.
- Wittenbrink, B., Judd, C. M., & Park, B. (1997). Evidence for racial prejudice at the implicit level and its relationship with questionnaire measures. *Journal of Personality and Social Psychology*, 72, 262-274.

Received October 20, 2005

Revision accepted February 3, 2006