Published in final edited form as:

J Exp Psychol Hum Percept Perform. 2015 April ; 41(2): 525–541. doi:10.1037/xhp0000039.

The Effect of Stereotype Threat on Performance of a Rhythmic Motor Skill

Meghan E. Huber,
Department of Bioengineering, Northeastern University

Allison E. Seitchik,
Department of Psychology, Northeastern University

Adam J. Brown,
Department of Psychology, Northeastern University

Dagmar Sternad, and
Departments of Biology, Electrical and Computer Engineering, and Physics, Northeastern University

Stephen G. Harkins
Department of Psychology, Northeastern University

Abstract

Many studies using cognitive tasks have found that stereotype threat, or concern about confirming a negative stereotype about one's group, debilitates performance. The few studies that documented similar effects on sensorimotor performance have used only relatively coarse measures to quantify performance. Three experiments tested the effect of stereotype threat on a rhythmic ball bouncing task, both at the novice and skilled level. Previous analysis of the task dynamics afforded more detailed quantification of the effect of threat on motor control. In this task, novices hit the ball with positive racket acceleration, indicative of unstable performance. With practice, they learn to stabilize error by changing their ball-racket impact from positive to negative acceleration. Results showed that for novices, stereotype threat potentiated hitting the ball with positive racket acceleration, leading to poorer performance of stigmatized females. However, when the threat manipulation was delivered after having acquired some skill, reflected by negative racket acceleration, the stigmatized females performed better. These findings are consistent with the mere effort account that argues that stereotype threat potentiates the most likely response on the given task. The study also demonstrates the value of identifying the control mechanisms through which stereotype threat has its effects on outcome measures.

Keywords

Stereotype Threat; Mere Effort; Rhythmic Movements; Motor Control; Social Threat

---

Correspondence regarding this article should be addressed to: Meghan E. Huber, 360 Huntington Avenue, 134 Mugar Life Sciences Building, Boston, MA 02115. mehuber@coe.neu.edu.

Stereotype threat refers to the concern that is experienced when one feels “at risk of confirming, as self-characteristic, a negative stereotype about one's group” (Steele & Aronson, 1995). A large number of studies have found that concern about confirming the relevant negative stereotype negatively impacts the performance of the stigmatized individuals (e.g., Aronson, Fried, & Good, 2001; Ben-Zeev, Fein, & Inzlicht, 2005; Blascovich, Spencer, Quinn, & Steele, 2001; Brown & Pinel, 2003; Davies, Spencer, Quinn, & Gerhardstein, 2002; Jamieson & Harkins, 2007; Johns, Schmader, & Martens, 2005; Schmader & Johns, 2003; Spencer, Steele, & Quinn, 1999; Steele & Aronson, 1995). A variety of cognitive tasks have been used to study these debilitating effects, including GRE verbal and quantitative problems (Steele & Aronson, 1995), tests of memory (Hess, Auman, Colcombe, & Rahal, 2003), GMAT problems (Quinn & Spencer, 2001), mental rotation problems (Martens, Johns, Greenberg, & Schimel, 2006), the Stroop Color-Word task (Jamieson & Harkins, 2011), and reading span tasks (Mazerolle, Regner, Morisset, Rigalleau, & Huguet, 2012).

In contrast, very few studies have examined the effect of stereotype threat on the performance of sensorimotor tasks. Three studies have examined golf putting (Stone, Lynch, Sjomeling, & Darley, 1999; Beilock, Jellison, Rydell, McConnell, & Carr, 2006; Stone & McWhinnie, 2008), one investigated soccer dribbling (Chalabaev, Sarrazin, Stone, & Cury, 2008), one driving in a simulator (Yeung & von Hippel, 2008), and one basketball free throw shooting and tennis serving (Hively & El-Alayli, 2014). Not only have few tasks been studied, but also performance on these tasks has been quantified with only relatively coarse measures. For example in the soccer dribbling task, Chalabaev et al. (2008) evaluated overall speed and found that females under stereotype threat completed the drill significantly more slowly than those without threat. While this finding supports the argument that the stereotype has a negative effect on sensorimotor performance, these results are mute about *how* stereotype threat affects motor performance. In soccer dribbling, slower performance can result either from being slow and cautious, but making minimal errors, or from moving fast, but making hasty mistakes that then need correcting. In golf putting, stigmatized individuals have been shown to be less accurate (Beilock et al., 2006) and require more strokes (Stone & McWhinnie, 2008) to putt a ball to a target hole. As in soccer, these outcome measures say little about how stereotype threat affects sensorimotor control: the increased number of strokes does not reveal whether more cautious behavior or larger errors led to specific changes in the golf swing without additional measures of sensorimotor control. Finally, a recent study by Hively and El-Alayli (2014) investigated the effect of stereotype threat on collegiate athletes in two different sensorimotor tasks, shooting a basketball and serving a tennis ball, by combining their performances with z-scores.

Performance on complex sensorimotor tasks is the result of a variety of motor, as well as cognitive, processes, and global performance measures alone cannot identify how stereotype threat may affect them. In fact, even if stereotype threat has little or no visible effect on the primary outcome measure, it may still affect the underlying control processes. For example, Chalabaev and colleagues (2013) found that, on a ballistic contraction task, ST did not affect the primary performance measure, maximum force production, but did influence the peak rate of force production, suggesting some effect on control processes. To investigate how stereotype threat affects sensorimotor control, it is necessary to derive measures that more directly reflect these control processes. For ball tasks such as golf putting, soccer dribbling, tennis serving, and basketball shooting overall performance is determined by the motor actions of the human as well as the dynamics of the ball. Extracting measures that reflect control requires an accurate model of the ball and the external environment.

Rhythmic Ball Bouncing Task

In the present work, we used an experimental task developed by Sternad and colleagues where the participants rhythmically bounced a ball with a racket (Schaal, Sternad, & Atkeson, 1996; Sternad, 1999, 2006; Sternad, Huber, & Kuznetsov, 2014). Participants manipulate a real table tennis racket to bounce a virtual ball to a target height, with movements restricted to the vertical direction (Figure 1). The deviation of the maximum ball height from the target height for each bounce served as the outcome measure or error that participants are instructed to minimize. Because the ball motion was simulated in a virtual environment, there were no uncontrolled aspects as would occur in a real-life version of the task; the ball dynamics were completely known. Knowing the exact physical model of the task allowed us to obtain underlying measures critical for control in addition to the typical performance measure, error.

To extract variables that delineate potential control strategies, a mathematical model of this task was developed and subsequently analyzed (Sternad, Duarte, Katsumata, & Schaal, 2000; Wei, Dijkstra, & Sternad, 2007; Dijkstra, Katsumata, de Rugy, & Sternad, 2004). The model consisted of a planar surface performing periodic vertical movements repeatedly impacting a ball. The model analysis of this nonlinear dynamical system showed that the system has dynamically stable solutions, meaning that small errors or perturbations die out by themselves. This approach is advantageous because the performer need not adapt his/her racket movements to every small deviation of the ball to maintain successful performance. Importantly though, dynamic stability is only achieved when the racket hits the ball during the decelerating portion of the racket's upward motion. Figure 2 illustrates this behavior for simple sinusoidal movement of the racket. Note that for this argument the racket movement does not need to be strictly sinusoidal, only periodic, as only the segment of the ball-racket contact matters. Hence, hitting the ball with negative acceleration is an efficient solution to this task, as small errors need not be corrected.

To make this property of the task more intuitive, consider the following: while hitting the ball with negative racket acceleration, balls that hit the racket earlier (such as after a lower ball amplitude in the preceding bounce) are hit with relatively higher velocity. This in turn leads to a higher ball amplitude on the following bounce, which is equivalent to an automatic correction of the previous low amplitude bounce. Conversely, balls that hit the racket later (such as after an amplitude overshoot) are hit with relatively lower velocity, which leads to a lower ball amplitude on the next bounce.

Previous research has shown that participants learn to exploit dynamic stability (e.g., Sternad, Duarte, Katsumata, & Schaal, 2001; Dijkstra et al., 2004; Sternad, 2006; Wei, Dijkstra, & Sternad, 2008; Huber & Sternad, under review). Novices initially hit the ball during the accelerating portion of the racket's upward motion to impart energy to the ball in the upward direction. In contrast to the dynamic stability achieved with ball-racket impacts during the decelerating racket trajectory, hitting the ball with an accelerating trajectory produces unstable performance and errors amplify from one bounce to the next (Figure 2B,C). To compensate for such unstable performance, the novice participants can actively correct for errors by adjusting their racket trajectory to propel the ball either higher or lower than the previous bounce, based on visual information about the error (de Rugy, Wei, Müller, & Sternad, 2003; Wei et al., 2007; Siegler, Bazile, & Warren, 2013). However, with practice, participants learn to hit the ball with negative acceleration, reducing the necessity for active correction of errors (Wei et al., 2008). Thus, this analysis identifies a control variable, racket acceleration, through which stereotype threat may have its effect on the outcome variable, error.

Theories on Stereotype Threat

The effect of stereotype threat on task performance has been conceptualized in two lines of work, one emphasizing performance on cognitive tasks (e.g., Schmader, Johns & Forbes, 2008) and the other performance on sensorimotor tasks (e.g., Beilock, Jellison, Rydell, McConnell & Carr, 2006). Schmader et al. (2008) have proposed a model that is primarily aimed at accounting for debilitation on cognitive tasks. This model incorporates cognition, affect, and motivation and how they impact performance through their effects on working memory efficiency. Performance is debilitated because concern about fulfilling the stereotype occupies working memory resources that could be used for task performance.

This model also incorporates a separate pathway for performance on sensorimotor tasks, thereby including results by Beilock and her colleagues (e.g., Beilock et al., 2006). Beilock argues that stereotype threat and performance pressure result in a two-pronged effect whereby people not only worry about the situation (thereby depleting working memory as in Schmader et al., 2008), but also explicitly monitor their performance in order to ensure optimal performance (e.g., DeCaro, Thomas, Albert, & Beilock, 2011). Beilock has focused on the performance of skilled athletes, and has found that pressure leads to the monitoring of performance, which disrupts well-learned, proceduralized behavioral sequences (e.g., Beilock, 2010; Beilock et al., 2006; Beilock & Carr, 2001).

Schmader et al.'s (2008) account incorporates motivation, but only indirectly, as it suggests that its effects manifest themselves by impacting efficiency of working memory. Beilock's account (e.g., Beilock et al., 2006) argues that threatened participants are motivated to minimize mistakes, but the cause of the performance debilitation is the resulting step-by-step focus on task execution. In the current work, we focus on the effect of stereotype threat on motivation and its direct effects on control and task performance by adopting Jamieson and Harkins's (2007) mere effort account.

Mere Effort: A Motivational Account

The mere effort account was suggested by Harkins's (2006) analysis of the effect of the potential for evaluation on performance. This account argues that the potential for evaluation motivates participants to want to perform well, which potentiates whatever response is prepotent, or most likely to be produced, on the given task. For example, on the Stroop Color-Word Task (Stroop, 1935), the prepotent response is to state the color word, as opposed to the correct response, naming the color in which the word is printed. If the prepotent response is incorrect and participants do not know, or lack the knowledge or time required for correction, performance is debilitated (Figure 3). However, if the prepotent response is correct, or if participants are able to recognize that their prepotent tendencies are incorrect and are given the opportunity to correct, performance will be facilitated. Harkins and his colleagues have found support for these predictions on the Remote Associates Task (Harkins, 2006), anagrams (McFall, Jamieson & Harkins, 2009, Experiment 1), the Stroop task (McFall et al., 2009, Experiments 2 and 3) and the antisaccade task (McFall et al., 2009, Experiment 4).

For example, Witte and Freund (2001) found that when solving anagrams, the initial tendency was to try consonants in the first position. McFall et al. (2009, Experiment 1) argued that subjecting participants to the potential for evaluation should potentiate this prepotent response. As a result, participants subject to evaluation should be better at solving anagrams of words that begin with consonants, but worse at solving anagrams that begin with vowels than participants not subject to evaluation. This is what they found. The same process (potentiation of the prepotent response) led to better performance in one case (words that began with consonants), but worse performance in the other (words that began with vowels).

Jamieson and Harkins (2007) argued that stereotype threat, like the potential for evaluation, motivates threatened participants to want to perform well. In this case, the motivation to counter the negative stereotype brings into play the same processes that are implicated in the evaluation-performance relationship. They have found support for this account, using the antisaccade task (Jamieson & Harkins, 2007), GRE quantitative problems (Jamieson & Harkins, 2009), and the Stroop (Jamieson & Harkins, 2011). The present research examines the effects of stereotype threat on sensorimotor performance and tests the predictions of the mere effort account.

The Present Research and Hypotheses

Based on previous research on the ball bouncing task, predictions can be made for the effect of stereotype threat on the outcome variable and potential control mechanisms, both for novices and experienced participants. As noted above, in the virtual ball bouncing task, the initial tendency of novices is to hit the ball with positive acceleration. Adopting the perspective of the mere effort account, this behavior represents the prepotent or most likely response for novice performers. Hence, the mere effort account argues that stereotype threat should potentiate positive racket acceleration at impact in novices, leading to worse performance than controls. In contrast, experienced performers have learned to hit the ball with negative acceleration to exploit dynamic stability. Hence, hitting with negative acceleration is the prepotent response in experienced participants, who should therefore perform better than control participants under stereotype threat.

For novices, Schmader et al.'s (2008) working memory account would not appear to provide an a priori basis for predicting the effect of threat on acceleration or for acceleration's effect on error. For experienced participants, Beilock's explicit monitoring account predicts that stereotype threat would lead to debilitation, not facilitation, as threat leads them to think about enacting behaviors that have been proceduralized and run effectively without the contribution of working memory.

In the present research, we tested the predictions of the mere effort account for male and female novice and experienced performers using the rhythmic ball bouncing task. All participants were told that performance on this visuo-spatial task was highly related to math ability, and that gender differences either had (stereotype threat) or had not (no stereotype threat) been found in performance on this task. This stereotype threat was intended to stigmatize the female participants, and thus potentiate their prepotent response.

In Experiments 1 and 2, novice participants performed 25 trials of the ball bouncing task; half of the male and female participants received the stereotype threat manipulation before starting the task. For these experiments on novice performers, we hypothesized that the prepotent response, hitting the ball with positive racket acceleration, would be potentiated in the stigmatized females. As a result, novice females under stereotype threat would perform with larger errors than control novices. We expected no performance differences between novice males under the same stereotype threat conditions.

In Experiment 3, female participants performed 12 trials of the ball bouncing task. At this point, a stereotype threat manipulation was implemented after which the participants performed an additional 12 trials. We hypothesized that in the initial 12 trials, the participants would learn to exploit dynamic stability such that in the second set of 12 trials the prepotent response would be to hit with negative racket acceleration. This response would then be potentiated in the females under stereotype threat, resulting in lower errors than control females.

Experiment 1

In Experiment 1, male and female participants were asked to perform 25 trials of the ball bouncing task, each lasting 40 sec. Prior to performing, participants were randomly assigned to either the stereotype threat or to the no-stereotype threat condition. We hypothesized that stereotype threat would potentiate the prepotent response, hitting the ball with positive racket acceleration, making it more difficult for threatened females to hit the ball with negative acceleration. As performance would not utilize dynamic stability, errors from prior bounces would not be stabilized, and the threatened females should perform more poorly than non-threatened females. We further hypothesized that this effect would be absent in males.

Method

Participants—Seventy-two undergraduate students (36 males and 36 females) from Northeastern University participated in the experiment in exchange for partial fulfillment of a course requirement. None had any prior experience with the specific task.

Task—In the experimental task, the participant stood 2 m in front of rear projection screen (2.43 m × 2.43 m) holding a real table tennis racket in his or her dominant hand (Figure 1). A light rigid rod with two hinge joints was attached to the racket surface and ran through a wheel whose rotation was registered by an optical encoder at a sampling rate of ~500 Hz. While the joints allowed the racket to move and tilt with minimal friction in all three dimensions, the encoder only measured the vertical displacement of the racket. The position of the virtual racket on the screen, represented by a horizontal red line (0.2 m × 0.02 m), was controlled by the measured position of the real racket. The vertical position of the virtual ball, represented by a white filled circle (0.02 m radius), was determined using ballistic flight and inelastic instantaneous impact equations. Based on these equations, the maximum ball height of each bounce was determined by the ball velocity, racket velocity, and racket height at ball-racket impact. To simulate the haptic sensation of a real ball hitting the racket, a mechanical brake was attached to the rod that was activated at ball-racket impact of each bounce. The participant was instructed to rhythmically bounce the ball to a target line, represented by a yellow horizontal line (1.0 m × 0.02 m) extending from the left edge to the middle of the screen, positioned 1 m above the minimum racket position.

Procedure—Each trial began with the ball appearing on the target line at the left side of the screen and rolling horizontally to the center of the screen. Upon reaching the center, the ball dropped vertically from the target line to the virtual racket. The participant was instructed by Experimenter 1 (male) to continuously bounce the ball to the exact vertical height of the target line for the duration of the trial. Each trial lasted 40 seconds and consisted of approximately 60-80 bounces.

All participants were given one practice trial for familiarization under the supervision of Experimenter 1 (male). After this practice trial, the experimenter informed the participants of the following: “The task you are about to complete is a test of visuo-spatial capacity. Performance on this task is closely linked to math ability. As you may know, there has been some controversy about whether there are gender differences in math and spatial ability. Previous research has demonstrated that gender differences exist on some of these tasks, but not on others. In our lab, we examine performance on both kinds of tasks.” Participants were then randomly assigned to a stereotype threat (ST) or no stereotype threat (N-ST) condition. ST participants were informed that the task had been shown to produce gender differences, whereas N-ST participants were told that it had not. This verbal instruction has been shown to produce stereotype threat effects in previous research (e.g., Brown & Pinel, 2003; Keller & Dauenheimer, 2003; O'Brien & Crandall, 2003; Spencer et al., 1999). At this point, Experimenter 1 excused himself after introducing Experimenter 2 (female), who was blind to each participant's condition. Experimenter 2 supervised each participant as she or he performed 25 trials (with a brief rest after 12 trials).

Following completion of 25 trials, the participants filled out a brief questionnaire for evaluating the effectiveness of the stereotype threat manipulation. The questionnaire asked: to what extent are there gender differences in performance on this task (1 = no gender differences and 11 = gender differences), who do you believe performs better on this task? (1 = males perform better, 6 = males and females perform the same, and 11 = females perform better), and to what extent is performance on this task related to mathematical ability (1 = not at all and 11 = closely related).

Data reduction and dependent measures—Each bounce in the trial was defined as the event between two consecutive ball-racket impacts, or two consecutive ball position minima (Figure 4). Task performance was characterized by the median error of bounces in each trial. Error was defined as the unsigned difference between the target height and the maximum ball amplitude for each bounce in the trial. Median racket acceleration at impact of each trial was used as a criterion whether participants performed with dynamically stable solutions. To obtain acceleration of the racket trajectory, the racket position was resampled at a constant frequency of 1 kHz and double-differentiated using a second-order Savitzky-Golay filter with a window size of .05 s (Savitzky & Golay, 1964). Racket acceleration at impact was defined as the racket acceleration 20 ms prior to the first minimum of ball position at each bounce. The interval of 20 ms was chosen to avoid capturing any artifacts due to the activation of the mechanical brake.

To identify the presence of active error corrections by the participants, it was necessary to tease apart the automatic error stabilization due to dynamic stability from such additional active error-based corrections. To do so, we followed the approach detailed in Wei et al. (2008) that quantified the self-stabilizing properties of the map with the autocorrelation function. The autocorrelation of successive ball release velocities at impact determined how fast errors dissipate, or exponentially decay, across impacts. The correlation values of the function relate to the time constant of exponential error decay over bounces. Positive autocorrelation values correspond to exponential decay, with faster decay indicated by smaller positive values. However a simple autocorrelation analysis of bounce-to-bounce fluctuations was not sufficient as errors also decline as a function of dynamic stability. To distinguish between error decreases due to dynamic stability and active corrections Wei et al. (2008) assessed the error dynamics in a ball bouncing model with an added stochastic component. As this model did not include any error correction, the autocorrelation values quantified the amount of error dissipation due to dynamic stability. Wei et al. (2008) concluded that if the lag-1 autocorrelation values obtained from participant performance were more negative than those generated by the stochastic model, then the participant applied active error corrections to decrease the error faster than with error stabilization alone.

Comparison of the lag-1 autocorrelation value from the stochastic model with the autocorrelation in participant performance could therefore quantify the relative amount of active error correction. Using the same parameters from the virtual task in the stochastic model rendered a lag-1 autocorrelation value of 0.24. Therefore a lag-1 autocorrelation value from participant performance below 0.24 indicated the presence of active error correction in a given trial.

Statistical analysis of performance—All performance measures were analyzed in 2 (Threat) × 2 (Gender) × 25 (Trials) analyses of variance (ANOVAs) with threat and gender as between-subjects factors and trials as a within-subjects factor. The Greenhouse-Geisser correction factor was applied to the within-subject effects (Kirk, 1995). Prior to these analyses, all dependent measures were checked for deviations from normal distributions using W-tests (Shapiro and Wilk, 1965). Medians of absolute error and racket acceleration were used because their deviations within each trial were not strictly Gaussian.

Results

Perception of stereotype threat—The measures of perception of threat obtained from the questionnaires were analyzed in 2 (Threat) × 2 (Gender) between-subjects ANOVAs. Participants in the ST condition reported that gender differences existed in this task to a greater extent ($M = 7.11$, $SD = 2.63$) than in the N-ST condition ($M = 3.43$, $SD = 2.81$), $F(1, 68) = 34.86$, $p < .001$, $d = 1.43$. Neither the main effect for gender, nor the interaction was significant, *F*s < 1, *p*s > .30.

ST participants also tended to report that males performed better on the task ($M = 4.35$, $SD = 1.96$) compared to N-ST participants ($M = 4.97$, $SD = 1.38$), $F(1, 68) = 2.65$, $p = .11$, $d = .39$. Neither the main effect for gender, nor the interaction was significant, *F*s < 1, *p*s > .40. Although the threat main effect was only marginal, the same manipulation check was used in each of the three experiments reported in this work. Taken together across the three experiments, the effect was highly reliable with a combined $z = 3.98$, $p < .0001$, $d = .64$. The effect for the gender difference question was also highly reliable across the three experiments, combined $z = 6.78$, $p < .00001$, $d = 1.41$.

Analysis of the question asking how related performance on the task was to mathematical ability revealed no significant differences, *F*s < 1, *p*s > .70. The grand mean on this measure ($M = 6.96$, $SD = 2.62$) was significantly different from the midpoint of the scale (6), $t(71) = 3.10$, $p < .01$, $d = .37$.

Performance error—The primary performance measure on the racket task was the median absolute error in each trial for each participant. Given that the target height determined the number of bounces for a given trial duration, each trial had on average 50 bounces. The 2 (Threat) × 2 (Gender) × 25 (Trial) ANOVA on the median error revealed a significant main effect for trial, $F(24, 1632) = 15.21$, $p < .0001$, reflecting the fact that the magnitude of error dropped over the course of the 25 trials from 26 cm ($SD = 23$ cm) in trial 1 to 8 cm ($SD = 10$ cm) in trial 25 (Figure 5A). The analysis also revealed a marginal main effect for threat, $F(1, 68) = 2.74$, $p < .11$, $d = .40$. Participants in the ST condition tended to exhibit larger errors across the 25 trials ($M = 13$ cm, $SD = 14$ cm) than participants in the N-ST condition ($M = 10$ cm, $SD = 11$ cm). Neither the main effect for gender nor the Gender × Threat interaction reached significance, *F*s < 1, *p*s > .70. The Trial × Gender, $F(24, 1632) = 1.61$, $p = .15$ and the three-way interaction, $F < 1$, $p > .50$, were also not significant.

Racket acceleration at impact—Racket acceleration at impact served as the measure of dynamic stability in task performance. As observed in previous studies, participants initially hit the ball with positive racket accelerations until they learned to exploit the dynamically stable solution that produced automatic stabilization for small errors. Consistent with these results, the median racket acceleration at impact again revealed a significant trial effect, $F(24, 1632) = 38.10$, $p < .0001$, indicating that the racket acceleration values changed from positive in trial 1 (*M* = 3.34 m/s^2, *SD* = 3.81 m/s^2) to negative values in trial 25 (*M* = -1.16 m/s^2, *SD* = 2.90 m/s^2) (Figure 5B). Analysis of this measure also revealed a main effect for gender, *F*(1, 68) = 4.47, $p < .05$, *d* = .51. Overall, females tended to hit the ball with more negative racket acceleration (*M* = -0.85 m/s^2, *SD* = 3.28 m/s^2) than males (*M* = 0.32 m/s^2, *SD* = 3.11 m/s^2). Neither the threat main effect nor the Gender × Threat interaction was significant, *F*s < 1, *p*s > .40. The interaction for Gender × Trial was significant, *F*(24, 1632) = 2.43, $p < .02$, as females hit the ball with greater positive racket acceleration in trial 1 (*M* = 3.78 m/s^2, *SD* = 3.58 m/s^2) than males (*M* = 2.89 m/s^2, *SD* = 4.02 m/s^2), *F*(1, 1632) = 4.26, $p < .05$, but ended the experiment with lower acceleration values, showing more improvement (e.g., trial 25: *F*(1, 1632) = 10.91, $p < .05$; females: *M* = - 1.88 m/s^2, *SD* = 2.57 m/s^2; males: *M* = -0.44 m/s^2, *SD* = 3.07 m/s^2). Neither the Trial × Threat, $F < 1$, $p > .80$, nor the three-way interaction was significant, *F*(24, 1632) = 1.08, $p > .35$.

Lag-1 autocorrelation of release ball velocities at impact—Lag-1 autocorrelation of release ball velocities at successive impacts measured the presence of error stabilization and active error correction. A *t*-test revealed that the mean lag-1 autocorrelation value of all participants was positive (*M* = .13, *SD* = .11), consistent with results from previous analysis (Wei et al., 2008). Importantly, the autocorrelation values for both males and females were significantly lower than 0.24, the value obtained from stabilization in the stochastic ball bouncing model without error correction. This suggested that additional control by participants reduced the correlations and suggested active error corrections. ANOVA revealed that lag-1 autocorrelation significantly decreased across trials, *F*(24, 1632) = 6.83, $p < .0001$, from .32 (*SD* = .29) in trial 1 to .07 (*SD* = .20) in trial 25 (Figure 5C). Further, the main effect for gender was significant, *F*(1, 68) = 8.53, $p < .05$, *d* = .71. While the mean autocorrelation values for males and females were both significantly less than 0.24, males used significantly more error correction (*M* = .09, *SD* = .23) than females (*M* = .17, *SD* = .28). There was no main effect for threat, $F < 1$, $p > .65$, nor was there a Gender × Threat interaction, $F < 1$, $p > .45$. The Trial × Threat, $F < 1$, $p > .90$, the Trial × Gender, *F*(24, 1632) = 1.31, $p > .18$, and the three-way interaction, *F*(24, 1632) = 1.05, $p > .39$, were also not significant.

Discussion

All participants learned to perform the task across practice as expected from previous results (Schaal et al., 1996; de Rugy et al., 2003; Wei et al., 2008). Both males and females showed similar improvements in the primary performance measure, median error. A marginal main effect for threat was seen as participants in the stereotype threat conditions tended to perform more poorly than their unthreatened counterparts. However, this effect was present in both males and females, counter to the hypotheses. Moreover, there was no hint of an effect of stereotype threat on racket acceleration at impact, the measure of dynamic stability.

One possible reason why we did not see the hypothesized Gender by Threat interaction is that we underestimated the effect of active error correction on task performance. Based on past research we expected active correction to play little, if any, role in performance (Wei, Dijkstra, & Sternad, 2007). Instead, over trials, we found decreases in both racket acceleration values and the measure of active error correction, the lag-1 autocorrelations of successive ball release velocities, suggesting that participants learned to employ both active correction and dynamic stability to improve their performance. The findings also suggest that males relied on active correction more than dynamic stability, whereas females relied more on dynamic stability than on active correction.

It is possible that the complex interplay of the two strategies accounts for the absence of the predicted stereotype threat effect. Ehrlenspiel, Wei and Sternad (2010) faced a similar problem in previous research using the same task. In their first experiment, participants learned to perform the ball bouncing task over the course of 32 trials on the first day. On the second day, the participants were randomly assigned to a high-stress or to a no-stress group. The high-stress participants were told that they had been entered in a competition with another participant. The participants then performed another 32 trials of the task. Ehrlenspiel et al. (2010) found that the participants subject to performance pressure improved more from day 1 to day 2 than participants in the control group. This finding is consistent with the mere effort account in that the participants learned to perform the task on day 1. As a result, the correct responses probably became more likely and, potentiated by the performance pressure, produced better performance in the competition condition. However, just as in our experiment, there were no differences in the measures of active control and dynamic stability. These differences were likely absent because participants may have adopted different approaches to error reduction, similar to our experiment.

In the ball bouncing task, it is not possible to identify one trial as generated by active correction and another by dynamic stability, as both strategies are simultaneously employed. Thus, in a second experiment, Ehrlenspiel et al. (2010) introduced perturbations to the flight of the ball on all trials of Day 2. These perturbations were too large to be self-correcting due to dynamic stability. This effectively minimized the contribution of dynamic stability to error reduction and made active error correction the only possible solution for reducing error. The results still showed improved performance in the high-stress condition, but there was also evidence that this group used active control more than the no-stress group. We conclude that this increase in error correction in the modified task was possible, because other potential compensatory mechanisms provided by dynamic stability were absent.

Thus, Experiment 2 introduced a manipulation that prevented participants from actively correcting the observed error in the previous bounce. We hypothesized that removing this compensatory mechanism would better reveal the effect of stereotype threat on task performance.

Experiment 2

Similar to Experiment 1, male and female participants were asked to perform 25 trials of the ball bouncing task. Again, participants were randomly assigned to either the stereotype threat or the no-stereotype threat condition prior to performing the task. In this experiment, however, unbeknownst to the participants, we applied a time-dependent manipulation to the racket velocity at ball-racket impacts as described in Huber and Sternad (2014, under review). Instead of using the actual racket velocity at ball-racket impact to determine ball position for the virtual display, we used the racket velocity 25 samples (50 ms) prior to impact. This time shift of the racket velocity altered the mapping between the perceived error and the corrective action, making actions that were normally successful in correcting an error no longer effective. If the racket impacted the ball in the accelerating portion of the trajectory, the manipulation caused errors to propagate faster and performance became unstable. Because hitting the ball with positive acceleration led to higher errors that could not be actively corrected, the only way to successfully perform the task was to exploit dynamic stability by hitting the ball with negative racket acceleration.

Huber and Sternad (2014, under review) showed that there was no significant difference in error between participants who performed the task with and without the time-dependent manipulation, indicating that the manipulation did not change the difficulty of the task. Furthermore, they found that even in this manipulated task, the median racket acceleration at impact was initially positive. This indicated that the prepotent response in the task with the time shift was the same as under previous conditions. Thus, we maintained the hypotheses from Experiment 1: stereotype threat potentiates the response of hitting with positive racket acceleration. As a result, the threatened females should perform more poorly and larger errors were expected. In contrast, threatened males should be unaffected by the threat. We expected to see this result more clearly than in Experiment 1, because it was necessary to hit the ball with negative racket acceleration to achieve error stabilization.

Methods

Participants—Sixty-nine Northeastern University undergraduate students (32 males and 37 females) participated in the experiment in exchange for partial fulfillment of a course requirement. None had any prior experience with the specific task.

Task and procedure—The task and procedure was identical to that described in Experiment 1, including all manipulations and questionnaire items. However, the calculation and simulation of the ball trajectory did not use the veridical racket velocity at ball-racket impact, but the racket velocity 25 samples (50 ms) prior to impact.

Data reduction and dependent measures—All data analysis and reduction measures were identical to Experiment 1 with two exceptions: 1) In the perturbed case, active error correction could no longer be determined and was excluded from the data analysis in this experiment; and 2) we needed to eliminate short portions of the trials. Due to the manipulation, the ball occasionally became unstable and “stuck” on the racket as the participant moved the racket up and down. During these “stuck bounces,” the ball did not reach the target line and the errors and other dependent measures were very large. Including these values in the average estimates of performance would have significantly skewed these measures, without adding to our understanding of how participants performed the task. Consequently, these uncharacteristic bounces were excluded from each trial before calculating the dependent measures. To identify these events, an investigator who was blind to the experimental condition defined a threshold distance of 0.25m between the maximum ball position and maximum racket position; bounces that were below this threshold were eliminated. Participants with a very high number of stuck bounces across trials were omitted from further analysis, as it was impossible to accurately assess their performance. The criterion for elimination of participants was when the mean duration of stuck bounces across trials was longer than two standard deviations away from the overall participant mean.

Statistical analyses—The performance measures were analyzed in 2 (Threat) × 2 (Gender) × 25 (Trials) ANOVAs, with threat and gender as between-subjects factors and trials as a within-subjects factor.

Results

Perception of stereotype threat—The perception of threat assessed in questionnaires was analyzed in 2 (Threat) × 2 (Gender) between-subjects ANOVAs. Participants in the ST condition reported that gender differences existed in this task to a greater extent ($M = 6.79$, $SD = 2.78$) than N-ST participants ($M = 3.97$, $SD = 2.93$), $F(1, 65) = 16.29$, $p < .001$, $d = 1.00$. Participants in the ST condition also reported that males performed better on the task to a greater extent ($M = 4.27$, $SD = 1.75$) than N-ST participants ($M = 5.29$, $SD = 1.25$), $F(1, 65) = 7.48$, $p < .01$, $d = .69$. Neither the gender main effect nor the Gender × Threat interaction was significant, *F*s < 1, *p*s > .80.

Once again, analysis of the question asking how related performance on the task was to mathematical ability revealed no significant differences, *F*s < 1, *p*s > .40. The grand mean on this measure ($M = 7.46$, $SD = 2.10$) was significantly different from the midpoint of the scale (6), $t\ (68) = 5.78$, $p < .0001$, $d = .70$. This pattern of findings shows that the stereotype threat manipulation was successfully implemented.

Failed performance—Applying the exclusion criterion based on mean duration of stuck bounces, four participants (two males and two females) were excluded from further analysis. They exhibited an average of 13.5 to 20 s per 40-s trial, where the ball was close to the racket and never achieved regular rhythmic behavior. This substantially exceeded the overall mean of 2.94 s per trial. An analysis performed on the mean duration of stuck bounces across trials for the remaining 65 participants revealed no reliable group differences, *p*s > .16. The analysis yielded a significant main effect for trial, $F(24, 1464) = 14.40$, $p < .0001$. Across trials, the amount of time spent in this behavior dropped from an average of 6.87 s in trial 1 to 0.73 s in trial 25.

Performance error—Analysis of the performance error revealed a significant change across trials, $F(24, 1464) = 29.42$, $p < .0001$, reflecting the fact that the participants' performances improved from trial 1 ($M = 21$ cm, $SD = 9$ cm) to trial 25 ($M = 10$ cm, $SD = 7$ cm). Participants reached an asymptote at about trial 18 (Figure 6A). This analysis also revealed a reliable Gender × Threat interaction, $F(1, 61) = 4.24$, $p < .05$, $d = .53$. A planned contrast showed that the error for ST females was greater ($M = 17$ cm, $SD = 8$ cm) than for N-ST females ($M = 12$ cm, $SD = 6$ cm), $F(1, 61) = 7.35$, $p < .01$, $d = .69$. In contrast, the performance of ST males ($M = 11$ cm, $SD = 6$ cm) and N-ST males ($M = 12$ cm, $SD = 7$ cm) did not differ, $p > .50$ (Figure 6B). The main effect for gender, $F(1, 61) = 6.32$, $p < .05$, $d = .64$, and the marginal threat main effect, $F(1, 61) = 2.65$, $p = .11$, $d = .42$, must be interpreted in the context of the two-way interaction, suggesting that all participants learned the task, but the threatened females started out and remained at a lower skill level than the others. The

Gender × Trial, Threat × Trial, and three-way interaction were all nonsignificant, *F*s < 1, *p*s > .45.

Racket acceleration at impact—Analysis of the median racket acceleration at impact revealed a significant trial effect, $F(24, 1464) = 34.67$, $p < .0001$, which reflected the fact that the participants' racket accelerations at impact went from positive in trial 1 ($M = 5.02$ m/s^2, $SD = 4.75$ m/s^2) to negative in trial 25 ($M = -3.24$ m, $SD = 2.59$ m), reaching an asymptote at about trial 20 (Figure 6C). This analysis also produced a Gender × Threat interaction, $F(1, 61) = 8.39$, $p < .01$, $d = .74$. The average racket acceleration at impact across the 25 trials for threatened females was positive ($M = .54$ m, $SD = 5.23$ m/s^2), whereas it was negative for N-ST females ($M = -1.67$ m/s^2, $SD = 3.88$ m/s^2), $F(1, 61) = 7.28$, $p < .01$, $d = .69$ (Figure 6D). While both threatened and unthreatened females learned to hit with negative acceleration by the end of the experiment, threatened females hit with less negative racket acceleration throughout, as shown in Figure 6D. The fact that they hit with positive acceleration for approximately the first ten trials resulted in a positive mean overall. Racket acceleration at impact for ST males ($M = -1.68$ m/s^2, $SD = 3.51$ m/s^2) did not differ from N-ST males ($M = -0.41$ m/s^2, $SD = 4.28$ m/s^2), $p > .15$. The main effects for threat and gender, the two-way interactions, and the three-way interaction were all nonsignificant, *F*s < 1, *p*s > .40.

Discussion

As in Experiment 1, the participants' performance improved over the course of the 25 trials, which was paralleled by a decrease in racket acceleration at impact towards more negative values. Like Experiment 1, the performance of threatened females was significantly worse than that of non-threatened females. Unlike in Experiment 1, however, threatened females were less successful in exploiting dynamic stability than non-threatened females, consistent with the main hypothesis. In fact, the average racket acceleration at impact across all 25 trials was positive for threatened females, whereas it was negative for non-threatened females. For males, there was no difference in error or acceleration as a function of threat. The marginal debilitation effect found in the male/threat condition in Experiment 1 did not recur.

By applying the time-shift manipulation, we were able to see that threatened females had more difficulty performing the task than non-threatened females. Moreover, we were able to see that this difficulty stemmed from the fact that, at the outset, hitting the ball with positive racket acceleration was the prepotent response, and that threat potentiated this response. These findings are consistent with the mere effort account in that threat debilitates performance when the prepotent response is incorrect.

The mere effort account also predicts that when threat potentiates a prepotent response that is correct, performance is improved. This means that if the participants had already learned to exploit dynamic stability, the response of hitting the ball with negative racket acceleration would be prepotent, and should be potentiated by threat. As a result, threatened females would be expected to perform better, not worse, than non-threatened females. This hypothesis was tested in Experiment 3.

Experiment 3

In this experiment, we recruited only female participants, who performed 24 trials of the ball bouncing task with the same manipulation introduced in Experiment 2. However, the first block of 12 trials was performed without any threat manipulation. Examination of the error results in Experiment 2 had suggested that by trial 12 participants had learned to hit the ball with negative acceleration, and thereby acquired a certain level of expertise. At this point, the females were randomly assigned to a threat or a no-threat condition, using the same manipulation as that employed in Experiment 2. To the extent that the prepotent response was now to hit the ball with negative acceleration and exploit dynamic stability, we hypothesized that threat would potentiate this correct response, leading to a better performance by threatened females than their non- threatened counterparts. Thus, unlike Experiment 2, females subject to stereotype threat were hypothesized to perform better, not worse, than non-threatened females.

Methods

Participants—Thirty-five Northeastern University undergraduate females participated in the experiment in exchange for partial fulfillment of a course requirement. None had any prior experience with the specific task.

Task and procedure—The task and procedure were identical to that described in Experiment 2, including all manipulations and questionnaire items. The only difference was that the threat manipulation was implemented after 12 trials instead of at trial 1 and the experiment continued for another 12 trials. At the outset, a male experimenter described the ball bouncing procedure followed by a practice trial. At this point, he left the room and a female experimenter collected the data for the first 12 trials. She then indicated that she needed to take a brief break, and the male experimenter returned. He then implemented the stereotype threat manipulation under the guise of providing some background information on the research, while they waited for the other experimenter to return. The female experimenter then returned and was blind to the experimental condition as she collected the data for the last 12 trials.

Data reduction and dependent measures—The method of data reduction, participant exclusion criteria, and dependent measures were identical to those described in Experiment 2.

Statistical analysis—Experiment 3 was conducted in two 12-trial blocks, separated by the stereotype threat manipulation that was given after the first block. For each of the performance measures, we conducted a 2 (Threat) × 12 (Trials) ANOVA, one for Block 1 and one for Block 2.

Results

Perception of stereotype threat—The perception of threat measures were analyzed in a one-way ANOVA with Threat as the independent variable. Females in the ST condition reported that gender differences existed to a greater extent ($M = 6.06$, $SD = 3.19$) than N-ST females (M = 2.82, SD = 1.93), $F(1, 33) = 12.48$, $p < .01$, d = 1.23. Females in the ST condition also reported that males performed better on the task to a greater extent (M = 4.11, SD = 1.94) than N-ST participants (M = 5.65, SD = 1.27), $F(1, 33) = 7.59$, $p < .01$, d = .96.

Analysis of the question asking how related performance on the task was to mathematical ability revealed no significant differences, $p > .50$. The grand mean on this measure (M = 7.43, SD = 2.51) was significantly different from the midpoint of the scale (6), t (34) = 3.36, $p < .01$, d = .57. Once again, this pattern of findings showed that stereotype threat was successfully implemented.

Failed performance—First, the intervals with stuck bounces were determined in both blocks. On the basis of performance in Block 1, one participant was eliminated (average of 19.08 s of stuck bounces per trial vs. overall M = 6.14 s per trial). For the remaining 34 participants, the mean duration of stuck bounces across trials in Block 1 revealed a reliable trial effect, $F(11, 352) = 18.77$, $p < .0001$. The duration of stuck bounces dropped from an average of 15.71 s in trial 1 to 2.79 s in trial 12. There was no reliable threat effect, nor an interaction, ps > .30. In Block 2, the analysis of stuck bounces revealed no group difference, p = .19, no trial effect, p = .19, nor their interaction, $p > .40$. The participants spent 2.18 s in stuck bounces in the first trial of the second block and .49 s in the last.

Performance error—Analysis of the first block of 12 trials produced a reliable main effect for trial, $F(11, 352) = 20.36$, $p < .0001$. As shown in Figure 7A, the average median error dropped from 29 cm (SD = 10 cm) in trial 1 to 14 cm (SD = 8 cm) in trial 12. There was neither a threat effect, $F(1, 32) = 1.09$, $p > .30$ (Figure 7B), nor an interaction, $F(1, 32) = 1.03$, $p > .40$.

The second block of trials also showed a reliable trial effect, $F(11, 352) = 4.63$, $p < .001$. The errors dropped from a mean of 14 cm (SD = 7 cm) in trial 1 to 10 cm (SD = 5 cm) in trial 12 of Block 2. This analysis also revealed a reliable threat effect, $F(1, 32) = 5.48$, $p < .05$, d = .83. Females under stereotype threat performed better (M = 9 cm, SD = 4 cm) than their non-threatened counterparts (M = 13 cm, SD = 7 cm) (Figure 7B). The interaction of these variables was not significant, $F < 1$, $p > .60$.

Racket acceleration at impact—In Block 1, there was a reliable trial effect, $F(11, 352) = 10.32$, $p < .0001$. As shown in Figure 7C, median racket acceleration at impact dropped from a mean of 4.19 m/s^2 (SD = 5.86 m/s^2) in trial 1 to a mean of -1.26 m/s^2 (SD = 4.11 m/s^2) in trial 12. Neither the threat main effect, $F(1, 32) = 1.03$, $p > .30$ (Figure 7D), nor the interaction was significant, $F < 1$, $p > .65$

In Block 2, once again the trial main effect was reliable, $F(11, 352) = 5.69$, $p < .0001$, with the mean falling from -0.38 m/s^2 (SD = 5.07 m/s^2) in trial 1 to -3.08 m/s^2 (SD = 2.78 m/s^2) in trial 12. This analysis also revealed a significant stereotype threat main effect, $F(1, 32) = 4.70$, $p < .05$, d = .77. Threatened females hit the virtual ball with greater negative racket acceleration (M = -3.50 m/s^2, SD = 3.11 m/s^2) than non-threatened females (M = -1.81 m/s^2, SD = 3.07 m/s^2) (Figure 7D). The interaction was not significant, $F < 1$, $p > .50$

Discussion

The mere effort account suggests that social threat potentiates prepotent responses. If those responses are incorrect, threat debilitates performance as we found in Experiment 2. At the outset of the task the prepotent response was hitting the ball with positive acceleration, which was potentiated by threat, leading to poorer performance by threatened females than by non-threatened ones. However, in Experiment 3, females learned to hit the ball with negative acceleration prior to the implementation of the threat manipulation. As a result, the prepotent response was to hit with negative acceleration, and the potentiation of this response to exploit dynamic stability led to better performance by threatened participants than by non-threatened ones as hypothesized.

General Discussion

The effect of stereotype threat on the performance of stigmatized individuals has been studied in a variety of cognitive tasks. However, only a few studies have examined the effect of stereotype threat on sensorimotor performance, and in these studies performance has been quantified with only relatively coarse measures (e.g., distance from the ball to the target after a putt; the time needed to dribble through a slalom course). While such performance measures characterize overall performance, they do not capture more fine-grained aspects of the execution that may shed light on the mechanisms of control through which stereotype threat has its effects on performance.

Harkins (2006) previously raised a similar concern with regard to cognitive tasks, arguing that a comprehensive understanding of the task is necessary to gain insight into how social threat affects task performance. In fact, it was the in-depth analysis of the Remote Associates Task that led to the formulation of the mere effort account for stereotype threat effects (Jamieson and Harkins, 2007). The mere effort account suggests that participants do not simply fall victim to a process that negatively affects performance. Instead, it suggests that stereotyped participants actually intensify their efforts during task performance, but these efforts may be misdirected. The mere effort account argues that under stereotype threat the response that is most likely for the task is potentiated. If this response is correct, then performance will be improved. If incorrect, and the participant does not recognize this, or does not have the time or skill necessary for correction, performance will be debilitated. However, identifying the prepotent response at the control level, and its role in overall task performance requires a thorough understanding of the task demands.

In this study, we chose the sensorimotor skill of rhythmically bouncing a ball to a target height, because previous theoretical analysis of the task dynamics provided the foundation for teasing out the effect of threat on motor control (Schaal et al., 1996; Sternad et al., 2000). Whereas prior psychological studies used outcome measures, such as error, to draw inferences about the effect of stereotype threat, the ball bouncing task afforded more direct assessment of control-relevant variables. In the ball bouncing task, novice participants initially hit the ball with positive acceleration, and then learn to stabilize error by shifting ball-racket impact from positive to negative acceleration. Thus, the mere effort account predicts that performance should be debilitated for stigmatized novices, as their most likely (prepotent) response is to hit with positive acceleration. We tested this hypothesis in

Experiment 1. Counter to expectation, the results showed no significant difference between threatened and non-threatened females in the control-relevant variable, racket acceleration at impact. Further, threat had a tendency to debilitate the performance of both males and females in the stereotype threat condition. We attributed this finding to the fact that active error correction may have confounded results and complicated inferences about control and stereotype threat.

While exploitation of dynamic stability is a signature of expert performance, other control processes, such as active error correction can also be used to reduce error (de Rugy et al., 2003; Ronsse, Wei, & Sternad, 2010; Siegler, Bardy, & Warren, 2010; Siegler et al., 2013; Wei et al., 2007, 2008). In Experiment 1 we found that males and females used varying degrees of dynamic stability and active correction to achieve the same level of task performance. The complex interplay of these strategies to reduce error may account for the absence of the predicted effects. Had we only analyzed error, we would not have known that the interplay of the control mechanisms likely masked the effect of threat on error.

The results of Chalabaev et al. (2013) also illustrate the need to consider the effect of ST on other measures beyond overall performance. As their isometric force task was a simple one-step skill, the null effect of ST in the primary performance measure alone could lead to the interpretation that non-proceduralized tasks, which are not susceptible to explicit monitoring processes, are immune to the effects of stereotype threat. Instead, by considering a secondary measure, Chalabaev et al. (2013) revealed that ST did affect motor control even in the absence of explicit monitoring processes. These findings provide compelling evidence that outcome performance measures alone are not sufficient to shed light on the mechanisms through which stereotype threat affects motor performance.

Whereas Chalabaev et al. (2013) were able to capture the effect of ST by a second measure of performance, the complex ball bouncing task did not permit such a simple parsing of variables. Thus, we applied the time-shift manipulation with the goal of isolating the effect of ST on one of the control mechanisms. In earlier research on ball bouncing under stress, Ehrlenspiel et al. (2010) faced a similar problem in that they observed an effect of stress on error, but no effects on the mechanisms of control. In that research, as in Experiment 1, it is possible that some participants used active control, whereas others tended to rely on dynamic stability as a control mechanism. In a second experiment, Ehrlenspiel et al. therefore added random perturbations to the ball flight, which required the participants to use active error corrections and ruled out sole reliance on dynamic stability. This experimental variation led to the expected relationship between active correction and error.

Therefore, this study added a time-dependent manipulation to the racket velocity at impact in Experiments 2 and 3, which required participants to exploit dynamic stability to perform successfully. Using the modified task, Experiment 2 tested the initial hypothesis that performance by threatened novices would be degraded, because the prepotent response, hitting with positive acceleration, was potentiated. Consistent with this hypothesis, we found that the threatened novice females hit the ball with more positive racket accelerations and had higher errors than their unthreatened counterparts.

In Experiment 3, the stereotype threat manipulation was delivered after the females had practiced the task for 12 trials; having gained a moderate level of expertise, it was expected that the prepotent response would now be to hit with negative acceleration. We hypothesized that this response would be potentiated in experienced females stigmatized by stereotype threat, and thus facilitate their performance. Indeed, threatened females hit the ball with greater negative acceleration than non-threatened females, which was accompanied by smaller errors. A caveat is that we showed this overt effect only in the modified task. However, we infer that similar processes hold for the unmanipulated task, only masked by compensatory processes due to error correction.

The effect of stereotype threat on task performance has been conceptualized in two lines of work. With a focus on cognitive tasks, the model of Schmader et al. (2008) argues that the effect of stereotype threat on cognition, affect, and motivation combine to impair the efficiency of working memory. This model also has a pathway that incorporates findings of Beilock and colleagues on sensorimotor tasks (e.g., Beilock et al., 2006). Thus, it proposes that the performance of experienced participants is disrupted by stereotype threat, because the threat leads them to monitor processes that have been proceduralized (i.e., well-learned). A second line of theorizing (e.g., DeCaro et al., 2011) incorporates the same two processes, but has focused primarily on the effect of threat on the performance of sensorimotor tasks by experienced participants.

In accounting for the present results, one could argue that novices use working memory as they learn to perform the task, and that stereotype threat creates anxiety and self-doubt, which take up processing capacity, leading to the performance debilitation found in Experiment 2. For example, Stone et al. (1999) found that stereotype threat debilitated the performances of novice golfers. However, the working memory account provides no a priori basis for predicting the effect of threat on performance, nor does this account provide any basis for expecting racket acceleration to mediate the effect. In contrast, based on identification of the prepotent or most probable response, the mere effort account makes an a priori prediction for the effect that stereotype threat will have on the control mechanism and the terminal behavior.

For experienced participants, Beilock's explicit monitoring account predicts that threat would debilitate performance. For example, DeCaro et al. (2011) argue that in Beilock et al.'s (2006) golf putting task and in Chalabaev et al.'s (2008) soccer dribbling task expert performance was debilitated because stereotype threat led these participants to monitor their well-learned behaviors. However, in Ehrlenspiel et al. (2010) and in Experiment 3 of the current research, participants had learned to perform the task, which facilitated, not debilitated performance. Ehrlenspiel et al. (2010) suggested that this facilitated performance of their threatened participants may have been the result of the rhythmic nature of the task, which might make it less susceptible to the effects of explicit monitoring. The same caveat may also hold for the current experiments.

As mentioned in Ehrlenspiel et al. (2010), a previous brain imaging study (Schaal, Sternad, Osu, & Kawato, 2004), behavioral results (Ikegami, Hirashima, Taga, & Nozaki, 2010; Howard, Ingram, & Wolpert, 2011; Sternad et al., 2013), and modeling studies (Sternad, Dean, & Schaal, 2000; Ronsse, Sternad, & Lefevre, 2009) support the argument that different control strategies are used to perform rhythmic and discrete motor tasks. Hence, the influence of stereotype threat on these different control strategies could potentially lead to different performance effects in the discrete golf putting task (Beilock et al., 2006). However, this does not account for the debilitation observed in the soccer dribbling task (Chalabaev et al., 2008). The soccer dribbling task is a continuous, and possibly rhythmic task, similar to ball bouncing. The essential element in such continuous tasks is that errors continuously propagate from one ball contact to the next. In golf putting on the other hand, each putt is a new event with different initial conditions. While it is still unclear from these collective findings whether stereotype threat produces different effects in discrete and rhythmic tasks, it is an important consideration for future studies involving sensorimotor tasks.

Another potential reason for these divergent results could be that participants in these studies were not experienced in the specific experimental tasks. For example, in Beilock et al.'s (2006) research on golf putting, they pointed out that their “experts” were asked to putt the ball so that it stopped directly on the target. This specific experimental task is different from the task on which these participants are actually expert: putting the ball through the target into the hole. Beilock et al. (2006) acknowledged this fact, noting that experts found that the constraint of stopping the ball on the target made the task difficult. Similarly, the experimental task used in the soccer study by Stone et al. (2008) required participants to dribble with only the dominant foot, whereas they typically dribble with both feet. Thus, in neither case was the experimental task actually the task in which the participants were expert.

The detailed understanding of the ball bouncing task provides a theoretically grounded measure of expertise. In Experiments 1 and 2, the mean racket acceleration at impact in trial 1 was positive, assuring that the participants were indeed novices. In Experiment 3, participants were trained on the experimental task until they were experienced as indicated by negative racket acceleration at impact by trial 12 (Figure 7). Thus, the measure that determined the level of expertise was the same control mechanism that accounted for error performance on the specific ball bouncing task that we used. These considerations demonstrate how a fine-grained task analysis can contribute to a better understanding of how a psychological manipulation (e.g., stereotype threat) affects a behavioral measure.

These findings also highlight the fact that verbal task instructions to the participant can subtly affect motor learning and performance. Previously, different types of instructions have been shown to affect complex skill learning, most notably in terms of directing attentional focus (Perkins-Ceccato, Passmore, & Lee, 2003). These more subtle effects during uncontrolled, spontaneous interaction with the subject are generally ignored in motor control research. On the other hand, gender differences have been documented, particularly that boys have better motor and visuo-spatial abilities than girls (Müller & Sternad, 2004). More recently, modeling of brain activation networks has suggested a physiological basis for this gender difference (Ingalhalikar et al., 2013). In motor neuroscience these seemingly subtle effects of stereotype and gender on motor performance have been given little attention. Our results suggest that experimenters should take extra precautions to ensure that they do not indirectly elicit this stereotype in interactions with their participants to avoid confounding effects on experimental results.

Even if experimenters take such precautions in their interactions with their participants, it is still possible for subtle cues in the experimental setting itself to activate gender stereotypes. For example, research has shown that solo status (female tested by male experimenter along with other male participants) can produce threat effects (e.g., Inzlicht & Ben-Zeev, 2000; Schmader & Johns, 2003; Sekaquaptewa & Thompson, 2002). In fact, Stone and McWhinnie (2008) found that the gender of the experimenter by itself was sufficient to produce a threat effect in the context of a putting task. These findings suggest that investigators in motor control may want to consider assessing the beliefs of their participants about gender differences in a post-experimental questionnaire to ensure that gender stereotypes have not been inadvertently activated.

In conclusion, our research demonstrated the effect of stereotype threat on sensorimotor performance supporting a motivational explanation as formulated by the mere effort account. Our experimental results show that performance outcome measures alone are not sufficient to determine the mechanisms through which stereotype threat affects task performance. When using complex sensorimotor tasks to study the effects of ST, it is important to consider how it affects the underlying control mechanism(s), even if there is no detectable effect on overall performance. By demonstrating how subtle psychological influence can differentially affect motor learning, our findings also stress the need for care in verbal instructions when conducting motor learning experiments.

Acknowledgements

This research received funding from the U.S. Army Research Institute for the Behavioral and Social Sciences (Contract W5J9CQ-12-C-0046; PI: Stephen G. Harkins). The views, opinions, and/or findings contained in this report are those of the authors and shall not be construed as an official Department of the Army position, policy, or decision, unless so designated by other documents. DS was supported by The National Institutes of Health R01-HD045639, the American Heart Association 11SDG7270001, and the National Science Foundation NSF DMS-0928587. DS was also supported by a visiting scientist appointment and MH as a junior scientist at the Max-Planck Institute for Intelligent Systems in Tübingen, Germany. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the funding organizations.

Appendix

Model for the bouncing ball

Equations of motion for the ball. The vertical position of the virtual ball x_b between the *k*th and the *k+1*th racket-ball impact follows ballistic flight:

$$x_b(t) = x_b(t_k) + v_b^+ (t - t_k) - g/2(t - t_k)^2$$

where t_k is the time of the *k*th ball-racket impact, v_b^+ is the velocity of the ball just after impact, and g is the acceleration due to gravity (9.81 m/s^2). To determine the ball velocity just after impact v_b^+, an instantaneous impact is assumed as follows:

$$\alpha\left(v_b^-(t_k)-v_r^-(t_k)\right)=-\left(v_b^+(t_k)-v_r^+(t_k)\right)$$

where v_b and v_r are the racket and ball velocities just before (−) and after (+) impact, and the energy loss at the collision is governed by the coefficient of restitution. The mass of the racket is assumed to be much larger than the mass of the ball, such that the racket velocity does not change during impact:

$$v_r^-(t_k)=v_r^+(t_k)=v_r(t_k).$$

Thus the ball velocity just after impact was determined by:

$$v_b^+(t_k)=(1+\alpha)\,v_r(t_k)-\alpha v_b^-(t_k).$$

The racket and ball system can be modeled as a continuous dynamical system with sinusoidal racket motion. With this assumption, a discrete map can be derived based on two state variables, the ball velocity just after impact v_b^+ and the racket phase at impact θ_k. Local linear stability analysis of this discrete map identifies a period-1 attractor, when racket acceleration at impact a_r satisfies the inequality (Dijkstra et al., 2004; Schaal et al., 1996):

$$-2g\frac{(1+\alpha^2)}{(1+\alpha)^2}<a_r<0$$

Simulations of the ball bouncing map illustrate that when the impact occurs during negative racket acceleration of the upward racket swing, the ball exhibits stable period-1 behavior (Figure 2B). The map possesses other attractors besides the period-1 attractor, including a “sticking” behavior, where the ball follows the racket trajectory. This map exhibits “sticking” behavior when the ball-racket impact occurs during positive racket acceleration (Figure 2C).

Ball Bouncing Map with Time-Dependent Manipulation

Under normal conditions, as in Experiment 1, the ball velocity immediately after impact was determined by:

$$v_b^+(t_k)=(1+\alpha)\,v_r(t_k)-\alpha v_b^-(t_k).$$

In Experiments 2 and 3, the racket velocity at impact v_r was set equal to racket velocity 50 ms before the time of impact t_k. Thus the ball velocity just after impact was determined by:

$$v_b^+(t_k)=(1+\alpha)\,v_r(t_k-.05)-\alpha v_b^-(t_k).$$

As in the unperturbed map, the ball exhibits “sticking” behavior if the impact occurs during the positive racket accelerations (Figure 8A). During negative racket acceleration, however, initial impact phases that previously led to stable period-1 behavior in the unperturbed map now produce “sticking” behavior; only the more negative racket acceleration impact phases continue to produce stable period-1 behavior (Figure 8B). In fact, the time-dependent manipulation causes the domain of attraction for period-1 to shift by $.05\omega$ radians on the sinusoidal racket trajectory, where ω is the angular frequency (Figure 8C).

References

Aronson J, Fried CB, Good C. Reducing the effects of stereotype threat on African American college students by shaping theories of intelligence. Journal of Experimental Social Psychology. 2002; 38:113–125. http://dx.doi.org/10.1006/jesp.2001.1491.

Beilock SL, Jellison WA, Rydell RJ, McConnell AR, Carr TH. On the causal mechanisms of stereotype threat: Can skills that don't rely heavily on working memory still be threatened? Personality & Social Psychology Bulletin. 2006; 38:1059–1071. http://dx.doi.org/10.1177/0146167206288489. [PubMed: 16861310]

Ben-Zeev T, Fein S, Inzlicht M. Arousal and stereotype threat. Journal of Experimental Social Psychology. 2005; 41:174–181. http://dx.doi.org/10.1016/j.jesp.2003.11.007.

Blascovich J, Spencer SJ, Quinn D, Steele CM. African Americans and high blood pressure: The role of stereotype threat. Psychological Science. 2001; 12:225–229. http://dx.doi.org/10.1111/1467-9280.00340. [PubMed: 11437305]

Bosson JK, Haymovitz EL, Pinel EC. When saying and doing diverge: The effects of stereotype threat on self-reported versus non-verbal anxiety. Journal of Experimental Social Psychology. 2004; 40:247–255. http://dx.doi.org/10.1016/s0022-1031(03)00099-4.

Brown RP, Pinel EC. Stigma on my mind: Individual differences in the experience of stereotype threat. Journal of Experimental Social Psychology. 2003; 39:626–633. http://dx.doi.org/10.1016/s0022-1031(03)00039-8.

Chalabaev A, Brisswalter J, Radel R, Coombes SA, Easthope C, Clement-Guillotin C. Can stereotype threat affect motor performance in the absence of explicit monitoring processes?: Evidence using a strength task. Journal of Sport & Exercise Psychology. 2013; 35:211–215. [PubMed: 23535978]

Chalabaev A, Sarrazin P, Stone J, Cury F. Do achievement goals mediate stereotype threat?: An investigation on females' soccer performance. Journal of Sport & Exercise Psychology. 2008; 30:143–158. [PubMed: 18490787]

Davies PG, Spencer SJ, Quinn DM, Gerhardstein R. Consuming images: How television commercials that elicit stereotype threat can restrain women academically and professionally. Personality and Social Psychology Bulletin. 2002; 28:1615–1618. http://dx.doi.org/10.1177/014616702237644.

de Rugy A, Wei K, Müller H, Sternad D. Actively tracking ‘passive’ stability in a ball bouncing task. Brain Research. 2003; 982:64–78. http://dx.doi.org/10.1016/s0006-8993(03)02976-7. [PubMed: 12915241]

DeCaro MS, Thomas R, Albert NB, Beilock SL. Choking under pressure: Multiple routes to skill failure. Journal of Experimental Psychology: General. 2011; 140:390–406. http://dx.doi.org/10.1037/a0023466. [PubMed: 21574739]

Dijkstra TMH, Katsumata H, de Rugy A, Sternad D. The dialogue between data and model: Passive stability and relaxation behavior in a ball bouncing task. Nonlinear Studies. 2004; 11:319–344.

Ehrlenspiel F, Wei K, Sternad D. Open-loop, closed-loop and compensatory control: performance improvement under pressure in a rhythmic task. Experimental Brain Research. 2010; 4:729–741. http://dx.doi.org/10.1007/s00221-009-2087-8. [PubMed: 19943039]

Harkins SG. Mere effort as the mediator of the evaluation performance relationship. Journal of Personality and Social Psychology. 2006; 91:436–455. http://dx.doi.org/10.1037/0022-3514.91.3.436. [PubMed: 16938029]

Hess TM, Aumann C, Colcombe SJ, Rahhal TA. The impact of stereotype threat on age differences in memory performance. The Journals of Gerontology Series B: Psychological Sciences and Social Sciences. 2003; 58:P3–P11. http://dx.doi.org/10.1093/geronb/58.1.p3.

Hively K, El-Alayli A. “You throw like a girl:” The effect of stereotype threat on women's athletic performance and gender stereotypes. Psychology of Sport and Exercise. 2014; 15:48–55. http://dx.doi.org/10.1016/j.psychsport.2013.09.001.

Howard IS, Ingram JN, Wolpert DM. Separate representations of dynamics in rhythmic and discrete movements: evidence from motor learning. Journal of Neurophysiology. 2011; 105:1722–1731. http://dx.doi.org/10.1152/jn.00780.2010. [PubMed: 21273324]

Huber ME, Sternad D. Implicit guidance to dynamic stability in a rhythmic motor task. 2014 Manuscript under review.

Huber, ME.; Sternad, D. Implicit guidance to dynamic stability in a rhythmic ball bouncing.. Paper presented at the Northeast Bioengineering Conference (NEBEC); Boston, MA.. Apr. 2014

Ikegami T, Hirashima M, Taga G, Nozaki D. Asymmetric transfer of visuomotor learning between discrete and rhythmic movements. Journal of Neuroscience. 2010; 30:4515–4521. http://dx.doi.org/10.1523/jneurosci.3066-09.2010. [PubMed: 20335489]

Ingalhalikar, M.; Smith, A.; Parker, D.; Satterthwaite, TD.; Elliott, MA.; Ruparel, K.; Verma, R. Sex differences in the structural connectome of the human brain.. Proceedings of the National Academy of Sciences of the United States of America, Early Edition; 2013. p. 1-6.http://dx.doi.org/10.1073/pnas.1316909110

Inzlicht M, Ben-Zeev T. A threatening intellectual environment: Why females are susceptible to experiencing problem-solving deficits in the presence of males. Psychological Science. 2000; 11:365–371. http://dx.doi.org/10.1111/1467-9280.00272. [PubMed: 11228906]

Inzlicht M, McKay L, Aronson J. Stigma as ego depletion: How being the target of prejudice affects self-control. Psychological Science. 2006; 17:262–269. http://dx.doi.org/10.1111/j.1467-9280.2006.01695.x. [PubMed: 16507068]

Jamieson JP, Harkins SG. Mere effort and stereotype threat performance effects. Journal of Personality and Social Psychology. 2007; 93:544–564. http://dx.doi.org/10.1037/0022-3514.93.4.544. [PubMed: 17892331]

Jamieson JP, Harkins SG. The effect of stereotype threat on quantitative GRE problems: A mere effort interpretation. Personality and Social Psychology Bulletin. 2009; 35:1301–1314. http://dx.doi.org/10.1177/0146167209335165. [PubMed: 19407004]

Jamieson JP, Harkins SG. The intervening task method: Implications for measuring mediation. Personality and Social Psychology Bulletin. 2011; 37:652–661. http://dx.doi.org/10.1177/0146167211399776. [PubMed: 21393614]

Johns M, Schmader T, Martens A. Knowing is half the battle: Teaching stereotype threat as a means of improving women's math performance. Psychological Science. 2005; 16:175–179. http://dx.doi.org/10.1111/j.0956-7976.2005.00799.x. [PubMed: 15733195]

Keller J, Dauenheimer D. Stereotype threat in the classroom: Dejection mediates the disrupting threat effect on women's math performance. Personality and Social Psychology Bulletin. 2003; 29:371–381. http://dx.doi.org/10.1177/0146167202250218. [PubMed: 15273014]

Kirk, R. Experimental design. Brooks/Cole; Pacific Grove, CA: 1995.

Martens A, Johns M, Greenberg J, Schimel J. Combating stereotype threat: The effect of self-affirmation on women's intellectual performance. Journal of Experimental Social Psychology. 2006; 42:236–243. http://dx.doi.org/10.1016/j.jesp.2005.04.010.

Mazerolle M, Regner I, Morisset P, Rigalleau F, Huguet P. Stereotype threat strengthens automatic recall and undermines controlled processes in older adults. Psychological Science. 2012; 23:723–727. http://dx.doi.org/10.1177/0956797612437607. [PubMed: 22609539]

McFall S, Jamieson JP, Harkins SG. Testing the generalizability of the mere effort account of the evaluation–performance relationship. Journal of Personality and Social Psychology. 2009; 96:135–154. http://dx.doi.org/10.1037/a0012878. [PubMed: 19210071]

Müller H, Sternad D. Accuracy and variability in goal oriented movements: Decomposing gender differences in children. Journal of Human Kinetics. 2004; 12:31–50.

O'Brien LT, Crandall CS. Stereotype threat and arousal: Effects on women's math performance. Personality and Social Psychology Bulletin. 2003; 29:782–789. http://dx.doi.org/10.1177/0146167203029006010. [PubMed: 15189633]

Perkins-Ceccato N, Passmore SR, Lee TD. Effects of focus of attention depend on golfers' skill. Journal of Sports Science. 2003; 21:593–600. http://dx.doi.org/10.1080/0264041031000101980.

Quinn DM, Spencer SJ. The interference of stereotype threat with women's generation of mathematical problem-solving strategies. Journal of Social Issues. 2001; 57:55–71. http://dx.doi.org/10.1111/0022-4537.00201.

Ronsse R, Sternad D, Lefèvre P. A computational model for rhythmic and discrete movements in uni- and bimanual coordination. Neural Computation. 2009; 21:1335–1370. http://dx.doi.org/10.1162/neco.2008.03-08-720. [PubMed: 19018700]

Ronsse R, Wei K, Sternad D. Optimal control of a hybrid rhythmic-discrete task: the bouncing ball revisited. Journal of Neurophysiology. 2010; 103:2482–2493. [PubMed: 20130042]

Savitzky A, Golay MJE. Smoothing and differentiation of data by simplified least squares procedures. Analytical Chemistry. 1964; 36:1627–1639. http://dx.doi.org/10.1021/ac60214a047.

Schaal S, Sternad D, Atkeson CG. One-handed juggling: A dynamical approach to a rhythmic movement task. Journal of Motor Behavior. 1996; 28:165–183. http://dx.doi.org/10.1080/00222895.1996.9941743. [PubMed: 12529218]

Schaal S, Sternad D, Osu R, Kawato M. Rhythmic arm movement is not discrete. Nature Neuroscience. 2004; 7:1136–1143. http://dx.doi.org/10.1038/nn1322.

Schmader T, Johns M. Converging evidence that stereotype threat reduces working memory capacity. Journal of Personality and Social Psychology. 2003; 85:440–452. http://dx.doi.org/10.1037/0022-3514.85.3.440. [PubMed: 14498781]

Schmader T, Johns M, Forbes C. An integrated process model of stereotype threat effects on performance. Psychological Review. 2008; 115:336–356. http://dx.doi.org/10.1037/0033-295x.115.2.336. [PubMed: 18426293]

Sekaquaptewa D, Thompson M. The differential effects of solo status on members of high- and low-status groups. Personality and Social Psychology Bulletin. 2002; 28:694–707. http://dx.doi.org/10.1177/0146167202288013refer.

Shapiro SS, Wilk MB. An analysis of variance test for normality (complete samples). Biometrika. 1965; 52:591–611. http://dx.doi.org/10.2307/2333709.

Siegler IA, Bardy BG, Warren WH. Passive vs. active control of rhythmic ball bouncing: The role of visual information. Journal of Experimental Psychology: Human Perception and Performance. 2010; 36:729–750. http://dx.doi.org/10.1037/a0016462. [PubMed: 20515200]

Siegler IA, Bazile C, Warren WH. Mixed control for perception and action: Timing and error correction in rhythmic ball-bouncing. Experimental Brain Research. 2013; 226:603–615. http://dx.doi.org/10.1007/s00221-013-3475-7. [PubMed: 23515627]

Spencer SJ, Steele CM, Quinn DM. Stereotype threat and women's math performance. Journal of Experimental Social Psychology. 1999; 35:4–28. http://dx.doi.org/10.1006/jesp.1998.1373.

Steele CM, Aronson J. Stereotype threat and the intellectual test performance of African Americans. Journal of Personality and Social Psychology. 1995; 69:797–811. http://dx.doi.org/10.1037//0022-3514.69.5.797. [PubMed: 7473032]

Steele CM, Spencer SJ, Aronson J. Contending with group image: The psychology of stereotype and social identity threat. Advances in Experimental Social Psychology. 2002; 34:379–440. http://dx.doi.org/10.1016/s0065-2601(02)80009-0.

Sternad D. Juggling or bouncing balls: Parallels and differences in dynamic concepts and tools. International Journal of Sport Psychology. 1999; 30(4):462–489.

Sternad, D. Stability and variability in skilled rhythmic actions - a dynamical analysis of rhythmic ball bouncing.. In: Latash, ML.; Lestienne, F., editors. Motor control and learning. Springer; New York: 2006. p. 55-63.

Sternad D, Dean WJ, Schaal S. Interaction of rhythmic and discrete pattern generators in single-joint movements. Human Movement Science. 2000; 19:627–664. http://dx.doi.org/10.1016/s0167-9457(00)00028-2.

Sternad D, Duarte M, Katsumata H, Schaal S. Dynamics of a bouncing ball in human performance. Physical Review E. 2000; 63:011902–1-011902-8. http://dx.doi.org/10.1103/physreve.63.011902.

Sternad D, Duarte M, Katsumata H, Schaal S. Bouncing a ball: Tuning into dynamic stability. Journal of Experimental Psychology: Human Perception and Performance. 2001; 27:1163–1184. http://dx.doi.org/10.1037/0096-1523.27.5.1163. [PubMed: 11642701]

Sternad D, Marino H, Charles SK, Duarte M, Dipietro L, Hogan N. Transitions between discrete and rhythmic primitives in a unimanual task. Frontiers in Computational Neuroscience. 2013; 7(90):1–13. http://dx.doi.org/10.3389/fncom.2013.00090. [PubMed: 23355821]

Stone J, Lynch CI, Sjomeling M, Darley JM. Stereotype threat effects on Black and White athletic performance. Journal of Personality and Social Psychology. 1999; 77:1213–1227. http://dx.doi.org/10.1037//0022-3514.77.6.1213.

Stone J, McWhinnie C. Evidence that blatant versus subtle stereotype threat cues impact performance through dual processes. Journal of Experimental Social Psychology. 2008; 44:445–452. http://dx.doi.org/10.1016/j.jesp.2007.02.006.

Stroop JR. Studies of interference in serial verbal reactions. Journal of Experimental Psychology. 1935; 18:643–662. http://dx.doi.org/10.1037/h0054651.

Wei K, Dijkstra TMH, Sternad D. Passive stability and active control in a rhythmic task. Journal of Neurophysiology. 2007; 98:2633–2646. http://dx.doi.org/10.1152/jn.00742.2007. [PubMed: 17881482]

Wei K, Dijkstra TMH, Sternad D. Stability and variability: Indicator for passive stability in a rhythmic task. Journal of Neurophysiology. 2008; 99:3027–3041. http://dx.doi.org/10.1152/jn.01367.2007. [PubMed: 18353911]

Witte K, Freund J. Single-letter retrieval cues for anagram solution. Journal of General Psychology. 2001; 128:315–328. http://dx.doi.org/10.1080/00221300109598914. [PubMed: 11678360]

Yeung NCJ, von Hippel C. Stereotype threat increases the likelihood that female drivers in a simulator run over jaywalkers. Accident Analysis & Prevention. 2008; 40:667–674. http://dx.doi.org/10.1016/j.aap.2007.09.003. [PubMed: 18329419]

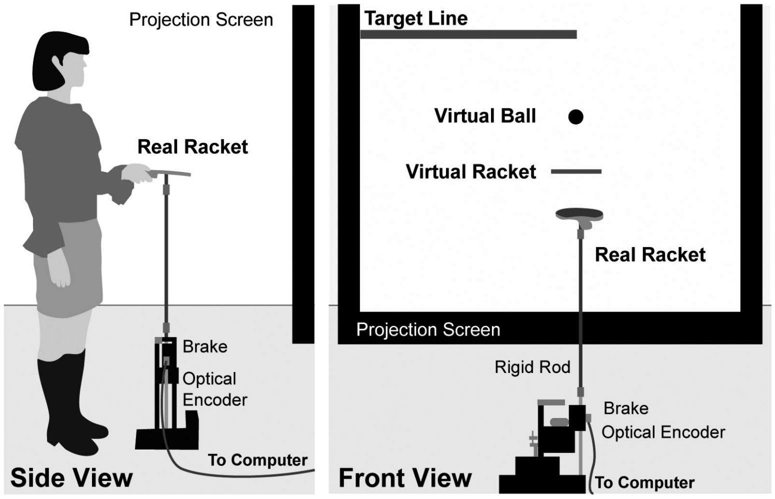

Figure 1.

Front and side view of the virtual experimental setup for ball bouncing. Participants were positioned in front of a screen and manipulated a real table tennis racket to rhythmically bounce a virtual ball to a target height in a 2D virtual environment.

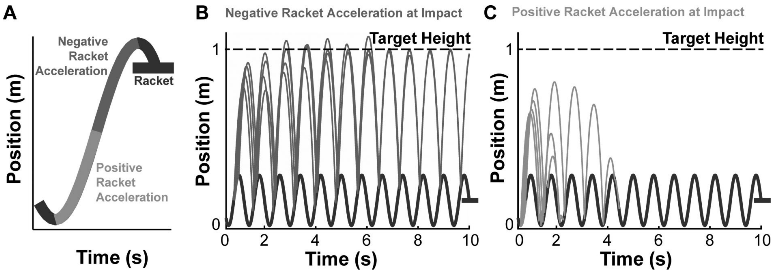

Figure 2.

Simulation of the ball-racket system. **A:** Assuming sinusoidal racket movement, the racket trajectory has a segment with positive acceleration followed by negative acceleration before its peak position. **B:** When the racket impacts the ball during the decelerating portion of the racket's upward motion, the ball-racket system is dynamically stable. Slightly different initial conditions all lead to the same stable ball amplitude without any changes in the racket trajectory. **C:** If the ball impacts the racket during the accelerating portion of the racket's upward motion, the system is unstable. Different initial conditions all lead to unstable behavior where the ball finally sticks to the racket. The only way to achieve and maintain a stable pattern is to correct for errors in the ball amplitude by a change in the racket trajectory.

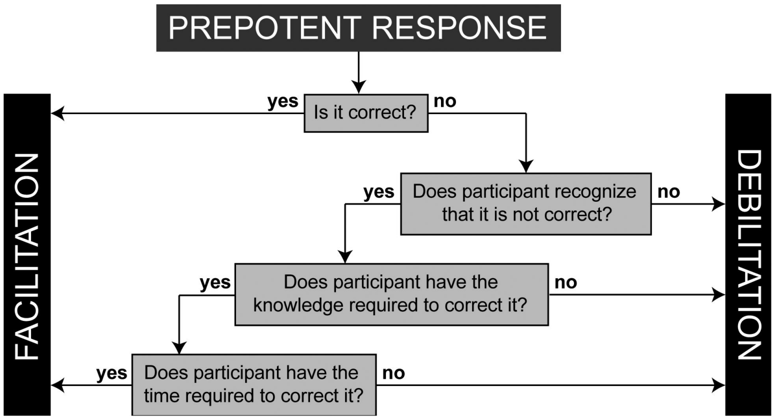

Figure 3.

Schematic diagram describing the mere effort account. If the prepotent response is correct, the mere effort account suggests that stereotype threat facilitates performance for stigmatized individuals. If the prepotent response is incorrect and participants do not know, performance is debilitated. However, if participants are able to recognize that their prepotent tendencies are incorrect and have the time to correct them, performance can be facilitated.

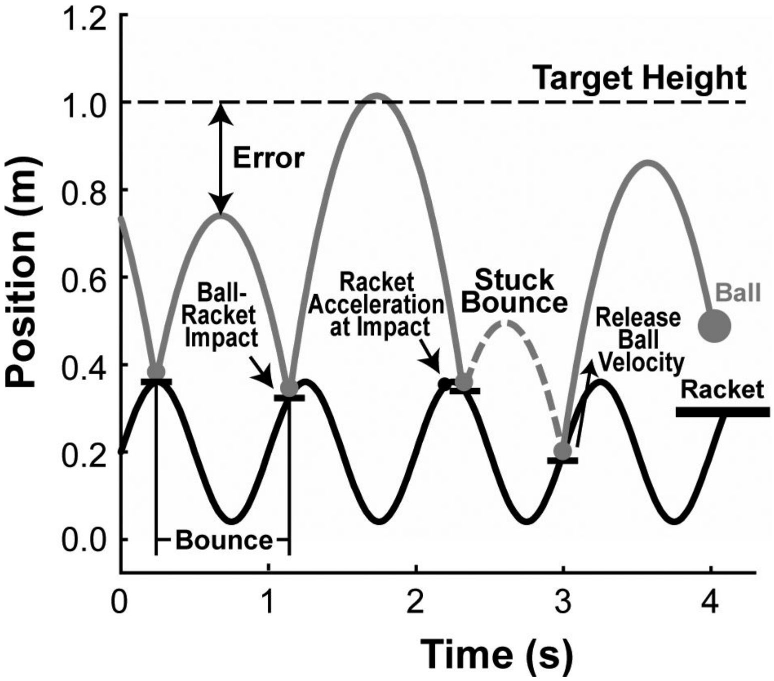

Figure 4.

Time series of racket (black) and ball (gray) trajectories illustrating the dependent measures. A bounce is the event between two consecutive ball-racket impacts. Error was defined as the unsigned difference between the target height and the maximum ball amplitude at each bounce. Racket acceleration at impact was defined as the racket acceleration 25 ms before the ball-racket impact of each bounce. Ball velocity at release was defined as the velocity of ball at the instantaneous ball-racket impact of each bounce. Bounces, where the difference between the ball and racket maximum positions was below 0.25 m, referred to as stuck bounces, were excluded from the data analysis in Experiments 2 and 3.

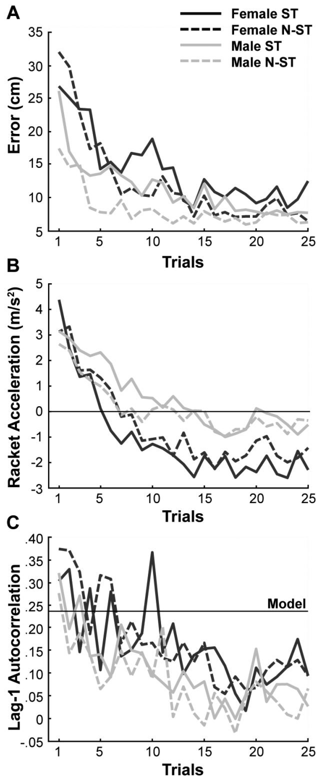

Figure 5.

Performance measures of all participants over practice in Experiment 1. **A:** The means of median of error of participants in all four experimental groups. **B:** Means of the median racket accelerations at impact for each trial across 25 trials. **C:** Means of lag-1 autocorrelation of release ball velocities at impact for each trial across 25 trials.

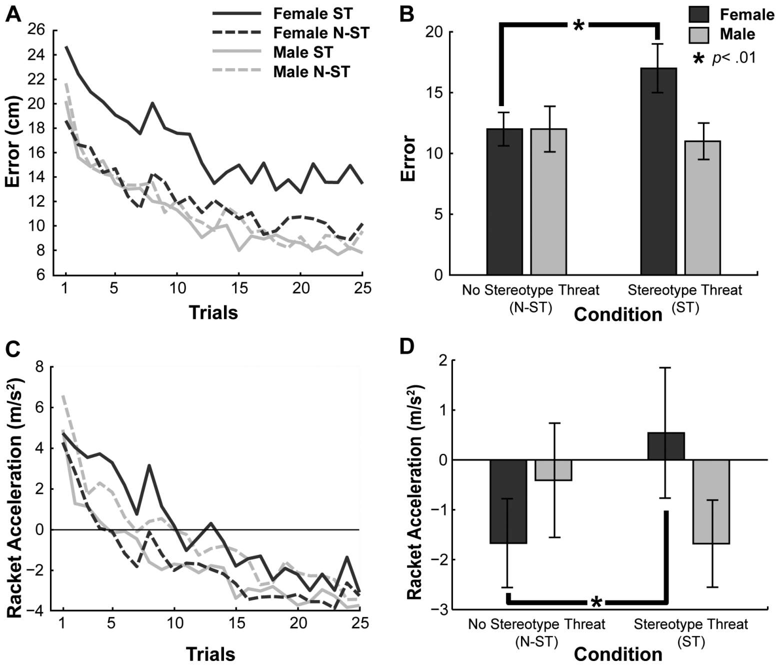

Figure 6.

Mean error and racket acceleration across practice and statistical comparisons in Experiment 2. **A:** Participant means of median errors across 25 trials. **B:** Statistical comparison of error between experimental groups. Error bars represent ± 1 standard error. **C:** Participant means of median racket accelerations across 25 trials. **D:** Statistical comparison of racket acceleration between experimental groups.

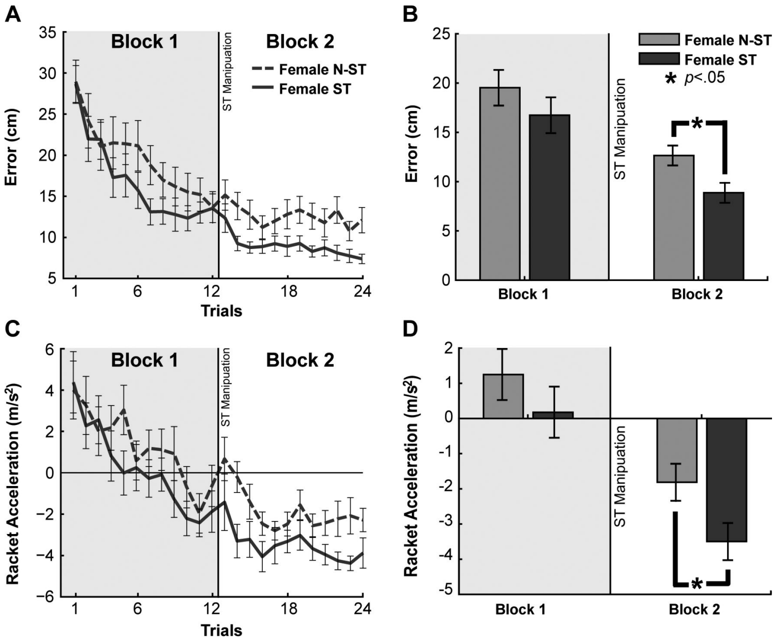

Figure 7.

Mean error and racket acceleration across practice and statistical comparisons in Experiment 3. Error bars represent ± 1 standard error of the mean. The Female ST group does not receive the stereotype threat manipulation until after trial 12, as marked with the vertical line. **A:** Participant means of median errors across 24 trials. **B:** Statistical comparison of error in Blocks 1 and 2 between experimental groups. **C:** Participant means of median racket accelerations across 24 trials. **D:** Statistical comparison of racket acceleration at impact in Blocks 1 and 2 between the two experimental groups.

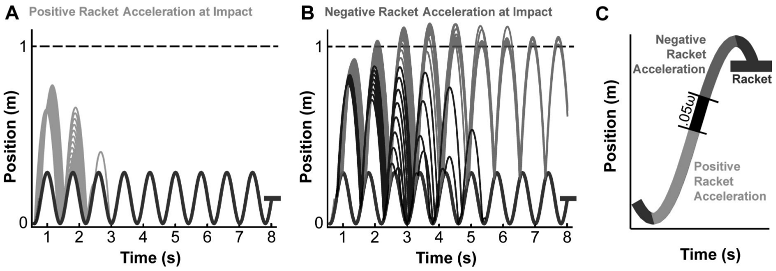

Figure 8.

Simulation of the ball-racket system with time-dependent manipulation. **A:** If the ball impacts the racket during the accelerating portion of the racket's upward motion, the system is unstable. **B, C:** When the racket impacts the ball during the decelerating portion of the racket's upward motion, the system is unstable if the racket phase is less than .05ω; otherwise, the perturbed ball-racket system is dynamically stable.