Article

Personality and Social Psychology Bulletin
2014, Vol. 40(3) 377–390

DOI: 10.1177/0146167213513475
pspb.sagepub.com

Stereotype Threat and Executive Functions: Which Functions Mediate Different Threat-Related Outcomes?

Robert J. Rydell[1], Katie J. Van Loo[1], and Kathryn L. Boucher[1]

Abstract

Stereotype threat research shows that women's math performance can be reduced by activating gender-based math stereotypes. Models of stereotype threat assert that threat reduces cognitive functioning, thereby accounting for its negative effects. This work provides a more detailed understanding of the cognitive processes through which stereotype threat leads women to underperform at math and to take risks, by examining which basic executive functions (inhibition, shifting, and updating) account for these outcomes. In Experiments 1 and 2, women under threat showed reduced inhibition, reduced updating, and reduced math performance compared with women in a control condition (or men); however, only updating accounted for women's poor math performance under threat. In Experiment 3, only updating accounted for stereotype threat's effect on women's math performance, whereas only inhibition accounted for the effect of threat on risk-taking, suggesting that distinct executive functions can account for different stereotype threat–related outcomes.

Keywords

stereotype threat, executive functions, math performance, risk taking

Received March 27, 2013; revision accepted October 25, 2013

There has been an explosion of research concerned with understanding stereotype threat, or individuals' worries about confirming negative stereotypes about their ingroup's ability with their own poor performance in the stereotyped domain (Steele, Spencer, & Aronson, 2002). Much of this work has examined the cognitive and affective mechanisms through which negative stereotypes hurt stereotyped individuals' performance in the stereotyped domain (Schmader & Beilock, 2012). Women who are reminded of the stereotype that "women are bad at math" perform more poorly at math than women who are not reminded of this stereotype (e.g., Beilock, Rydell, & McConnell, 2007). The most prominent theory about why this occurs claims that, because of negative thoughts, arousal, suppression, hypervigilance, and other related mechanisms, women under threat have reduced executive functioning, or "general purpose control mechanisms that modulate the operation of various cognitive subprocesses and thereby regulate the dynamics of human cognition" (Miyake et al., 2000, p. 50) and, thus, experience greater difficulty solving difficult math problems (Schmader, Johns, & Forbes, 2008). Reduced executive functioning under threat can also lead to self-regulatory impairment outside of the stereotyped domain ("spillover effects"; Inzlicht & Kang, 2010). Together, these findings demonstrate the role of general executive functioning in explaining various stereotype threat effects.

Despite the research linking general reductions in executive functioning to poor math performance among women experiencing stereotype threat, the more specific cognitive mechanisms through which stereotype threat impacts math performance or stereotype threat spillover effects have not been identified. Executive functioning as we refer to it here is a broader construct, including more cognitive processes and operations, than the construct of working memory, which is usually used to explain effects in the stereotype threat literature. Working memory is conceptualized as a limited resource of controlled executive attention processes in the stereotype threat literature (Schmader et al., 2008), but it is unclear if this general cognitive resource is responsible for the myriad of threat-related outcomes.[1] Stereotype threat researchers have often treated the interrelated, but independent, executive functions posited by more nuanced cognitive models of executive functioning (e.g., Miyake et al., 2000) interchangeably and as synonymous operationalizations of a general pool of cognitive resources, thus making it difficult to determine if different executive functions might be more

[1]Indiana University, Bloomington, USA

Corresponding Author:
Robert Rydell, Department of Psychological and Brain Sciences, Indiana University, 1101 E. 10th St., Bloomington, IN 47405, USA.
Email: rjrydell@indiana.edu

strongly implicated in accounting for different threat-related outcomes. We believe that adopting a more complex conceptualization of executive functioning can help refine our understanding of how stereotype threat has its impact. We argue that stereotype threat may affect several specific executive functions, which may account for different threat-related outcomes. Specifically, we are interested in examining which executive function(s) can account for the stereotype threat's impact on women's outcomes within the stereotyped domain (math performance) and outside of it (risk taking).

Stereotype Threat and Women's Math Performance

Most models explaining how stereotype threat hurts women's performance in math predict that negative performance stereotypes about their ingroup are inconsistent with the positive view that most women hold about themselves and their abilities (Nosek, Banaji, & Greenwald, 2002). People strive to maintain a positive view of the self despite negative self-relevant information (Baumeister, 1998). When women categorize themselves according to their gender, this positive view of the self is incompatible with negative stereotypes about women's math ability (Nosek et al., 2002; Rydell, McConnell, & Beilock, 2009), and the imbalance between positive self-views and negative self-relevant stereotypes leads women to experience a psychological state characterized by high levels of arousal, vigilance, worry, and rumination about confirming these negative stereotypes (Schmader et al., 2008). These responses to threat serve to reduce executive functioning, reducing math performance by usurping cognitive resources that could otherwise be devoted to math problems (e.g., Schmader & Johns, 2003).

Stereotype Threat "Spillover Effects"

While most work examining the underlying processes of stereotype threat has focused on women and math performance, research also suggests that executive functioning may be implicated in stereotype threat "spillover effects," or the impact of threat on outcomes outside of the stereotyped domain (e.g., Inzlicht & Kang, 2010; Inzlicht, McKay, & Aronson, 2006). Indeed, stereotype threat can affect important outcomes, such as overeating, aggression, persistence, risky choices, financial decisions, and verbal ability (Beilock et al., 2007; Carr & Steele, 2010; Inzlicht et al., 2006; Inzlicht & Kang, 2010). Many of these spillover effects are relevant to self-regulation (Inzlicht & Kang, 2010), and good executive functioning is important for successful self-regulation (see Hofmann, Schmeichel, & Baddeley, 2012).

Overeating, risk taking, and other outcomes experienced by women under threat may occur due to the same process as performance outcomes; temporarily reduced executive functioning impairs one's ability to muster the self-control to eliminate these behaviors. Here, we focus on one spillover effect. Inzlicht and Kang (2010) showed that people experiencing stereotype threat, relative to those who were not, were more likely to prefer a lottery with a large payout but a very low likelihood of winning over a lottery with a small payout but a high likelihood of winning (i.e., the small payout lottery is safer because it has a higher expected value). We examined this outcome because risk-taking is often detrimental, but it also seems that increased risk-taking in response to threat is potentially related to different cognitive functions than completing math problems. Whereas completing math problems utilizes complex computations and problem-solving strategies, determining whether or not to take a risk may rely on cognitive skills related to reducing the temptation of an alternative with a high payout to select a safer bet.

Executive Functions and Stereotype Threat

We adopted Miyake and colleagues' (2000) framework of executive functions to explore how stereotype threat might influence more specific executive functioning processes to impact various threat-related outcomes. Miyake et al. (2000) proposed that the executive functions of inhibition, shifting, and updating are the cognitive processes most important to effective executive functioning. Although these executive functions do not constitute short-term memory, they determine what is stored and how it is manipulated. While these functions are distinguishable and often predict different outcomes, they also exhibit unity because they are related at the level of latent variables and probably work in concert in many everyday situations (Miyake et al., 2000). In theory, and consistent with a unified component of executive functioning, it is possible that each of the executive functions could account for women's decreased math performance and increased risk-taking under threat. As explained below, however, we predict that the executive function of updating accounts for threat's detrimental effect on women's math performance, whereas inhibition accounts for women's increased risk preference when under threat. We do not expect shifting to account for either of these stereotype threat effects. But, it is important to stress that if threat does affect any of these executive functions, it is presumably because it leads to increased arousal, rumination, suppression, worry, and related processes, which then could negatively affect these executive functions (Schmader et al., 2008). Next, we define each executive function and describe the reasoning behind our predictions.

Updating

Updating captures the ability to utilize "attentional control to maintain relevant information (including task goals) in the face of interference, delete this information when it becomes irrelevant and replace it with new information"

(Friedman et al., 2006, p. 178). This executive function is the most closely related to conceptualizations of executive functioning in social psychology because "the essence of updating lies in the requirement to actively manipulate relevant information in working memory, rather than passively store information" (Miyake et al., 2000, p. 57). Updating, then, may tap into several important features underlying reduced performance on difficult math problems for women experiencing stereotype threat: online computation, renewing information about the math problem incrementally during problem solving, determining if incoming information is needed to solve the problem, deciding what information should be retained, and focusing on the goal of solving the problem.

Beilock et al. (2007) showed that because women become worried about confirming the negative math-related stereotypes, threat impairs their ability to complete the online computations necessary for difficult math problems, with the verbal ruminations utilizing the same resources. Research also shows that stereotype threat's detrimental effect on women's math performance is mediated by executive functioning (e.g., Schmader & Johns, 2003). In most experiments showing that executive functioning mediates the impact of threat on performance, executive functioning is measured by a complex "span" task (e.g., Conway et al., 2005), where people have to actively hold information in memory (e.g., several words to recall) while also intermittently engaging in online computations (e.g., counting the number of vowels in a sentence). Thus, the information to be retained and the online computations are vying for attentional resources. After completing an online computation, the output from that computation should be deleted so that new information can be acquired and actively maintained for later recall during the next online computation. Given the set-up of "span" tasks, the cognitive skills they measure seem quite similar to those involved in updating—the retention and manipulation of goal or task-relevant information and the deletion of goal or task irrelevant information (Schmiedek, Hildebrandt, Lövdén, Wilhelm, & Lindenberg, 2009). Because the skills involved in updating are likely impacted by stereotype threat and are also necessary to correctly solve difficult math problems, we expect updating to account for women's math performance under threat.

We do not expect that updating will account for the effect of stereotype threat on risk-taking. Although deciding to act in a risky manner, especially for a long time, may be partially driven by an inability to keep long-term goals actively represented in short-term memory (a part of updating; Hofmann et al., 2012), the ability to actively transform information, the ability to perform online computations, and the ability to avoid being distracted seem less relevant to risky choices. Whether or not people take risks seems more closely related to whether or not they act in line with, or can control, their impulses (e.g., Hofmann, Gschwendner, Wiers, Friese, & Schmitt, 2008).

Inhibition

Inhibition refers to the ability to deliberately suppress prepotent responses. That is, inhibition is the ability to exert mental effort to successfully control a relatively spontaneous response (see Miyake et al., 2000). This conceptualization of inhibition is defined very specifically in Miyake et al.'s (2000) model. Consistent with Miyake et al. (2000) and Hofmann et al. (2012), inhibition consists of an internally produced control response that can override a spontaneous cognition or behavior that is deemed "incorrect." Once a spontaneous response is stopped, it is necessary to identify or compute the correct response by using other cognitive resources.

When inhibition is couched in these terms, it is a limited but important construct. Women experiencing threat due to negative math-related stereotypes show reduced ability to inhibit their responses on a classic measure of inhibition (i.e., the Stroop [1935] task; Carr & Steele, 2010; Inzlicht et al., 2006). We propose that one consequence of reduced inhibition for women experiencing threat is that they will show greater risk preference. People who take risks often have difficulty regulating their behavior (i.e., resisting temptations) because they have difficulty controlling their impulses, urges, or automatic reactions (e.g., Baumeister & Vohs, 2004; McClure, Laibson, Lowenstein, & Cohen, 2004). Indeed, research shows that those with lower levels of executive functioning act in line with their automatically activated evaluations of a stimulus object (e.g., Hofmann et al., 2008), suggesting that their behavior will be more in line with their impulses (e.g., to make a risky choice with a high potential payout; see Masicampo & Baumeister, 2008). It is likely that reduced inhibition due to threat could also lead to people selecting the prepotent, risky response to the detriment of taking their time to weigh the pros and cons or calculate the expected values of different options. Therefore, we predict that reduced inhibition will account for increased risk taking under threat.

In math, women experiencing stereotype threat may have difficulty controlling their tendency to utilize easy-to-adopt but incorrect problem-solving strategies (Carr & Steele, 2009). Given this and the evidence that stereotype threat hurts inhibition (e.g., Inzlicht et al., 2006), inhibition may partially explain threat's effect on women's math performance. However, it seems unlikely that stopping a prepotent response would, in and of itself, be sufficient to account for poor math performance on *difficult* problems, which are the only types of problems negatively affected by stereotype threat (Beilock et al., 2007). Instead, more substantial controlled processing carried out by using other executive functions is likely needed to come up with a correct answer to a complicated math problem (e.g., a problem requiring multiple steps and symbol transformations). Therefore, we do not expect that inhibition will account for threat's impact on women's math performance.

Shifting

Shifting refers to people's ability to effectively transition between multiple tasks (Miyake et al., 2000). Women who are under threat may need to shift between the primary task (math performance) and other tasks (e.g., situational appraisals), with the reduced ability to shift between tasks resulting in poorer performance on the primary task. Although it is possible that problems with women's ability to shift their focus between math problems and the negative stereotype could hinder performance, shifting should not be detrimental to performance if it does not happen often and/or information related to the math problem can be actively maintained in short-term memory during shifting.

Research in the stereotype threat literature seems inconsistent with the view that difficulty shifting between tasks accounts for reduced math performance for women under threat. Stereotype threat can sometimes increase the number of problems attempted (Jamieson & Harkins, 2007) and performance on easy problems (O'Brien & Crandall, 2003). Neither of these results is consistent with people engaging in a large amount of shifting in response to threat or having difficulty actively maintaining information in memory when shifting under threat. Reduced shifting ability would presumably lead to lower levels of attention to the primary task, reducing the number of problems attempted. Similarly, reduced shifting ability, if it affected performance on easy problems, would hinder performance on easy problems. There is no research showing that stereotype threat reduces the number of math problems women attempt or that it hurts performance on easy problems. It seems unlikely that stereotype threat reduces shifting or that shifting would account for poorer math performance. Moreover, because shifting should not influence whether a risky alternative is enticing or whether people are able to control the temptation of a risky alternative, it should not account for the effect of threat on risk taking.

Overview of the Current Work

We examined the impact of stereotype threat on measures assessing the executive functions of inhibition, shifting, and updating, as well as the two different stereotype threat outcomes: math performance and risk taking. We expected that stereotype threat would negatively affect inhibition and updating. However, we expected that the negative effect of stereotype threat on updating would then account for reduced math performance among women, whereas we expected the negative effect of threat on inhibition to account only for their increased risk taking.

In Experiments 1 and 2, we establish that only updating accounts for women's math performance under stereotype threat. We expected updating to account for these effects because it necessitates utilizing attentional resources that are impaired by stereotype threat (e.g., Schmader et al., 2008) and the cognitive tasks governed by updating are seemingly necessary to solve difficult math problems. In Experiment 3, we included a risk-taking measure along with our math performance measure to demonstrate that specific executive functions underlie different threat-related outcomes. Risk taking is a domain where people might have an initial impulse that they would need to control to make a more rational decision. Many risky decision tasks involve foregoing a tempting option for an objectively superior option. Thus, people may have to inhibit the prepotent response to select the option with the large payout, and this ability may be hampered by stereotype threat. Therefore, we expected the executive function of inhibition to account for this effect of stereotype threat.

Experiment 1

In Experiment 1, we brought men and women into the lab and randomly assigned them to a control condition or a stereotype threat condition. They then completed a measure of each of the three executive functions and a math test. Consistent with past stereotype threat research, we expected that women's math performance would be lower in the stereotype threat condition than in the control condition. We did not expect men's math performance to differ as a function of threat. Importantly, we predicted that the executive function of updating, and not inhibition or shifting, would account for women's math performance as a function of threat.

Method

Participants. Male (n = 93) and female (n = 75) undergraduates participated for course credit. They were randomly assigned to a control condition or a stereotype threat condition.

Procedure. Participants were informed that our lab was examining cognitive skills. Participants first learned how to solve modular arithmetic (MA) problems. In MA, a math equation is presented, and participants determine if the outcome of the equation is "true" (i.e., the answer is a whole number) or "false" (i.e., the answer is not a whole number).

MA problems follow the format $a \equiv b$ (mod c), for example, $93 \equiv 49$ (mod 4). An equation is solved by subtracting b (49) from a (93) and then dividing the solution of $a - b$ (44) by c (4). If the solution of $a - b$ (44) divided by c (4) equals a whole number (11), the equation is "true." An equation that is "false," for example, $93 \equiv 47$ (mod 4), would result in a remainder when the solution of $a - b$ ($93 - 47 = 46$) is divided by c ($46 \div 4 = 11.5$). MA was learned by reading a detailed, step-by-step tutorial (Beilock et al., 2007).

Stereotype threat manipulation. Participants then received the manipulation of stereotype threat. Under the control condition, no reference was made to gender. Under the stereotype threat condition, participants were told that the research was investigating why women are generally worse at math than men (Beilock et al., 2007).

Executive functions tasks. Next, participants completed three tasks, one that measured each of the executive functions outlined by Miyake et al. (2000). The order of these tasks was randomized for each participant.[2] To measure inhibition, participants completed the Stroop (1935) task. In this task, participants were asked to indicate the color of the text (blue, red, or orange) for a color word (e.g., the word BLUE) that was presented at the center of the computer screen by pressing one of three response keys (marked by appropriately colored stickers). Participants were presented with a fixation cross in the center of the computer screen for 500 ms, followed by the presentation of the color word, which remained on the screen until participants indicated a response. On 48 incongruent trials, the color word was presented in a different colored text (e.g., BLUE was presented in red text). On 48 congruent trials, the color word and the color of the text were the same (e.g., BLUE printed in blue text). Because reading occurs more quickly than color identification, participants had to inhibit the response based on reading the color word and focus on the color of the text to correctly answer incongruent trials; however, this inhibition was not necessary when the color word and the color of the text were the same. In calculating participants' Stroop scores, we did not include incorrect trials or trials for which participants' responses were less than 300 ms or greater than 2000 ms. We then found the average reaction time for the remaining incongruent trials and the remaining congruent trials, subtracting the average reaction time for the incongruent trials from that of the congruent trials, with greater scores indicating greater inhibition. This scoring, which is opposite of normal scoring procedures, was done so that greater scores on all of our measures of executive functions would indicate higher functioning.

To measure updating, participants completed the letter-memory task (Morris & Jones, 1990). In this task, several letters were presented sequentially in the center of the computer monitor for 2500 ms each. Participants were asked to recall the last three letters presented for trials containing five, seven, or nine letters (there were 12 trials, 4 of each length) by typing them into a response box when prompted by the computer. Participants continually updated the letters in memory by verbally rehearsing the last three letters presented (dropping any letters out of verbal rehearsal that appeared 4 or more letters before). The number of correctly recalled letter triads divided by 12 served as our measure of updating. A greater proportion of correctly recalled letter triads indicated greater updating.

To measure shifting, participants completed the number-letter task. In this task, participants completed 108 trials in which they saw a horizontal line on the computer screen. Next, a cue (a box) was presented above the line (54 trials) or below the line (54 trials) for 150 ms. Then the box was filled with a letter and a number, which remained on the screen until participants made a response. If the box was presented above the line, participants indicated whether the number was odd or even by pressing one of two response keys ("c" for odd, "m" for even). If the box was presented below the line, participants indicated whether the letter was a consonant or a vowel using the response keys from the number judgment ("c" for vowel, "m" for consonant). When the box was presented in the same location on consecutive trials, it was a no-switch trial (53 trials). When the box changed location, it was a switch trial (53 trials). After eliminating incorrect response trials, a shifting score was calculated by subtracting the average response latency of switch trials from the average response latency of no-switch trials. Greater scores indicated greater shifting ability.

Math performance. Participants then completed 36 MA problems by indicating whether the equations presented were "true" (pressing the "t" key) or "false" (pressing the "f" key). The MA problems presented used larger numbers (i.e., 19-99) and required a borrow operation, making them difficult (Beilock et al., 2007). Accuracy was calculated by dividing the number of correctly answered MA problems by 36 (participants answered all problems). Greater scores indicated better math performance. The amount of time taken to complete each problem was averaged to measure math reaction time. Finally, participants indicated their gender.

Results

Correlations between executive functions. There was a significant positive correlation between the inhibition and updating measures, $r = .21$, $p = .006$. However, the correlations between the inhibition and shifting measures, $r = -.03$, $p = .74$, and the updating and shifting measures, $r = .04$, $p = .60$, were not significant.[3]

Executive functions tasks. Each of the executive functions tasks was submitted to a 2 (Gender: men, women) × 2 (Stereotype threat: control, stereotype threat) ANOVA (see Table 1). The results for the Stroop task showed a significant two-way interaction, $F(1, 164) = 5.73$, $p = .018$, $\eta_p^2 = .034$. Women showed poorer inhibition in the stereotype threat condition than in the control condition, $F(1, 164) = 7.95$, $p = .005$, $\eta_p^2 = .046$; however, men's level of inhibition did not differ as a function of threat, $F(1, 164) = 0.20$, $p = .653$, $\eta_p^2 = .001$. The results for the letter-memory task also showed a two-way interaction, $F(1, 164) = 15.24$, $p < .001$, $\eta_p^2 = .085$. Women showed poorer updating under the stereotype threat condition than under the control condition, $F(1, 164) = 20.89$, $p < .001$, $\eta_p^2 = .113$. Men's level of updating did not differ in response to the threat manipulation, $F(1, 164) = 0.57$, $p = .452$, $\eta_p^2 = .003$. The results for the number-letter task showed no significant effects, Fs < 1. Thus, shifting was not affected by gender, stereotype threat, or their interaction.

Table 1. Executive Function Tasks and Math Performance as a Function of Gender and Stereotype Threat in Experiment 1.

	Men		Women	
	Control	ST	Control	ST
Executive function tasks				
Inhibition task (Stroop)	$-109_{a,b}$	-103_{a}	-92_{a}	-131_{b}
Updating task (Letter memory)	0.89_{a}	0.91_{a}	0.92_{a}	0.78_{b}
Shifting task (Number letter)	-349_{a}	-384_{a}	-360_{a}	-349_{a}
Math performance				
Accuracy	0.88_{a}	0.90_{a}	0.90_{a}	0.80_{b}
Math reaction time	7.41_{a}	7.60_{a}	$8.40_{a,b}$	8.91_{b}

Note. ST = stereotype threat. The inhibition and shifting task scores are presented in milliseconds. Math reaction times are presented in seconds. Means in a row with different subscripts were significantly different at the $p < .05$ level.

Math performance. Accuracy and math reaction time were each submitted to a 2 (Gender) × 2 (Stereotype threat) ANOVA (see Table 1). The results for accuracy showed the expected two-way interaction, $F(1, 164) = 15.95$, $p < .001$, $\eta_p^2 = .089$. Women showed poorer accuracy under the stereotype threat condition than under the control condition, $F(1, 164) = 20.22$, $p < .001$, $\eta_p^2 = .110$. Men's accuracy did not differ as a function of threat, $F(1, 164) = 0.94$, $p = .334$, $\eta_p^2 = .006$. Math reaction times only showed a marginally significant main effect of gender, $F(1, 164) = 3.59$, $p = .060$, $\eta_p^2 = .021$. Women tended to take longer to complete an MA problem ($M = 8.66$ s) than men ($M = 7.50$ s). The lack of a two-way interaction of gender and stereotype threat on math reaction time, $F < 1$, indicates that the math performance results were not due to a speed-accuracy tradeoff.

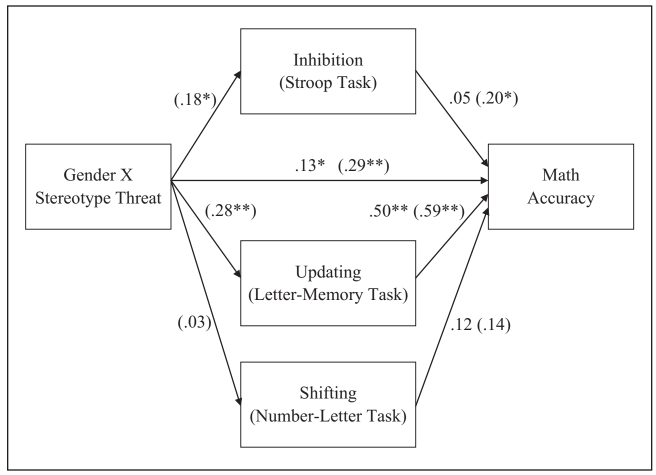

Figure 1. A path diagram showing the interaction of gender and stereotype threat on math accuracy as mediated through tasks assessing the executive functions of inhibition, updating, and switching in Experiment 1.
Note. All scores are standardized beta weights. The direct effects between variables are presented in parentheses.
$^*p < .05$. $^{**}p < .01$.

Mediational Analysis

To examine which executive functions accounted for the interaction of gender and stereotype threat on math performance, we used Preacher and Hayes's (2008) bias-corrected bootstrapping procedure for models with multiple mediators.[4] This model simultaneously examined the indirect effects whereby the interaction of gender and stereotype threat predicted each of the three executive functions, which in turn predicted math accuracy (see Figure 1).[5] In the model, the direct relation between the interaction of gender and stereotype threat still predicted math accuracy when executive functions scores were taken into account, but it was reduced. Although the interaction of gender and stereotype threat was related to both the Stroop (inhibition) task and the letter-memory (updating) task (see above), only the bias-corrected 95% confidence interval (CI) for the indirect effect of the letter-memory task on the relation between the interaction of gender and stereotype threat and math performance did not include 0 (+.007 to +.024). The bias-corrected 95% CIs for the indirect effect of the Stroop task and the number-letter task included 0 (−.004 to +.001 and −.003 to +.001, respectively). This indicates that only updating, and not inhibition or shifting, mediated the relation between the interaction of stereotype threat and gender and math performance.

Discussion

Experiment 1 showed that stereotype threat reduced women's math performance but had no effect on men's math performance. The effect of stereotype threat on women's math performance was partially accounted for by the executive function of updating; however, it was not explained by the executive functions of inhibition or shifting. These results provide a more detailed picture of the cognitive mechanisms that lead women experiencing threat to show poorer math performance. This experiment begins to examine the executive functions through which threat operates to reduce women's math performance, finding that difficulty updating seems to be integral to understanding why stereotype threat hurts women's math performance.

Experiment 2

In Experiment 2, we not only attempted to replicate the results of Experiment 1, but we also sought to strengthen our argument that updating accounts for women's reduced math performance as a function of stereotype threat. First, we changed the measure of math performance. MA may be very closely related to updating because it involves online manipulation of information, as well as performing mathematical operations and holding those outcomes in memory. In MA, people need to complete online calculations and keep the outcome from the first operation (subtraction) in memory because it is needed to perform the second operation (division). The similarities between MA and updating could inflate the relationship between these variables, perhaps accounting for the mediational results in Experiment 1. In Experiment 2, we used word problems from the GRE that are less likely to be as closely tied to updating as MA, and we gave participants paper to use during the math test, eliminating the need for all calculations performed to be computed online. If stereotype threat reduces women's math performance on these word problems and this reduction is accounted for by updating, and not inhibition or shifting, our claim that the reason why women's math performance is harmed by threat is because their ability to update is reduced would be strengthened.

Second, in many studies examining executive functions, several measures of each executive function are completed. Latent variables are then created from the common variance extracted from these measures, and these latent variables are used to predict the outcomes of interest (e.g., Friedman et al., 2006; Miyake et al., 2000). Because this procedure is lengthy (taking 90-120 min) and we wanted to maximize the chance that women were experiencing stereotype threat during math performance, we only used one measure of each executive function as in Experiment 1. However, we used different measures of updating and shifting to show that the effects generalize to different measures of these functions. If the results generalize across different measures of these executive functions, it reduces concerns that the results of Experiment 1 were due to aspects of the specific measures used.

We decided to retain our measure of inhibition in Experiment 2. Although women in Experiment 1 showed poorer inhibition in the stereotype threat condition than the control condition and inhibition was related to math performance, inhibition did not account for women's math performance. Given the overlap between updating and MA, inhibition might account for stereotype threat's effect on math performance when a different math test is used.

Method

Participants. We recruited only female participants for Experiment 2 because men showed no effects of stereotype threat in Experiment 1. Ninety female undergraduates, participating for course credit, were randomly assigned to a control condition or a stereotype threat condition.

Procedure. Participants were then given the same introduction and manipulation of stereotype threat that was used in Experiment 1.

Executive functions tasks. Next, participants completed three tasks, one that measured each of the executive functions. As in Experiment 1, the order of these tasks was randomized for each participant. As in Experiment 1, the Stroop task was used to measure inhibition in Experiment 2.

To measure updating, participants completed the keep track task (Miyake et al., 2000). In this task, participants completed 15 trials. Each trial consisted of 15 items (each presented for 1 s) that came from one of six categories (metals, countries, relatives, colors, animals, and distances). For each of the 15 items presented in a trial, the word was presented in the center of the screen and the category label was presented at the bottom of the screen. At the beginning of each trial, participants were given two, three, or four target categories (with the number of target categories presented dispersed evenly across the trials). The goal of the task was to recall the last item presented from the target categories. For example, if the target categories were relatives and animals, participants were told, "The target categories are RELATIVES and ANIMALS. Please remember the last RELATIVE and the last ANIMAL that you see listed." After participants saw the 15 items in a trial, they were presented with each of the target categories again and were prompted by a computer to recall, by typing their response into a box on the computer screen, the last item presented from each target category. In total, participants were given 45 items to recall. The greater number of correctly recalled items indicated a greater updating ability.

To measure shifting, participants completed the color–shape task. In this task, participants indicated either the color or the shape of a stimulus across 108 trials. In each trial, a cue (an "S" if they were to identify the shape on the trial and a "C" if they were to identify the color on the trial) was presented in the top half of the computer screen for 150 ms. Then, they saw both the cue (on the top half of the screen) along with one of four stimuli (on the bottom half of the screen; green circle, green triangle, red circle, red triangle) that remained on the screen until participants made a response. When cued with an "S," participants indicated whether the shape of the stimulus was a triangle or a circle by pressing one of two response keys ("f" for triangle, "k" for circle). When cued with a "C," participants indicated whether the color of the stimulus was red or green using the same response keys from the shape judgment ("f" for red, "k" for green). When the cue was the same on consecutive trials, it was a no-switch trial (53 trials). When the cue changed from the previous trial, it was a switch trial (53 trials). After eliminating trials for which participants responded incorrectly, a shifting score was calculated by subtracting the average response latency of switch trials from the average response latency of no-switch trials. Greater scores indicated greater shifting ability.

Math performance. Participants then had 10 min to complete 15 word problems on the computer that were taken from the GRE and used in past stereotype threat research (Schmader & Johns, 2003). Participants were presented with five potential answers for each problem and we also provided a "skip" button, which allowed them to move to the next problem without attempting to solve the problem presented.[6] The number of questions answered correctly (number correct), the number of questions attempted (number attempted), and the percentage of problems answered correctly (math accuracy = number correct divided by number attempted) were calculated and analyzed.

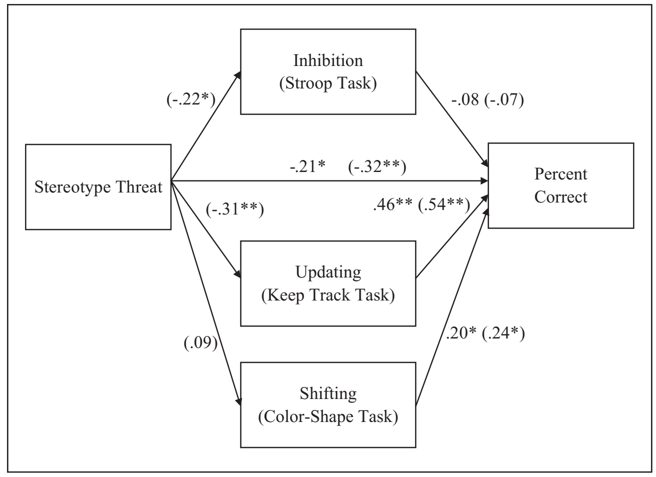

Figure 2. A path diagram showing the effect of stereotype threat on percent of math problems correct as mediated through tasks assessing the executive functions of inhibition, updating, and switching in Experiment 2.
Note. All scores are standardized beta weights. The direct effects between variables are presented in parentheses.
$^{*}p < .05$. $^{**}p < .01$.

Results

Correlations between the executive functions. The inhibition measure showed a marginally significant positive correlation with the updating measure, $r = .20$, $p = .063$, but was not correlated with the shifting measure, $r = .08$, $p = .45$. The updating measure was not correlated with the shifting measure, $r = .14$, $p = .17$.

Executive functions tasks. The effect of stereotype threat on each of the executive function tasks was examined using separate independent samples t tests. There was a significant effect of stereotype threat on the inhibition task, $t(88) = -2.09$, $p = .04$, $d = .44$. Women under threat showed poorer inhibition ($M = -141$ ms) than women in the control condition ($M = -107$ ms). The results for the updating task also showed an effect of stereotype threat, $t(88) = -3.07$, $p = .003$, $d = -.65$. Women showed poorer updating in the stereotype threat condition ($M = 27.32$) than in the control condition ($M = 30.30$). The stereotype threat manipulation did not affect performance on the shifting task, $t(88) = .81$, $p = .42$, $d = .17$ ($M_{\text{Control}} = -179$ ms, $M_{\text{Stereotype threat}} = -225$ ms).

Math performance. Women in the stereotype threat condition showed lower math accuracy ($M = 0.37$) than did women in the control condition ($M = 0.51$), $t(88) = -3.15$, $p = .002$, $d = -.66$. Women under threat also showed a lower number correct ($M = 3.36$) than did those under the control condition ($M = 5.15$), $t(88) = -5.15$, $p < .001$, $d = -1.09$. Women under the stereotype threat and control conditions attempted roughly the same number of math problems ($M = 10.36$, and $M = 11.35$, respectively), $t(88) = 1.25$, $p = .22$, $d = -.26$.

Mediational Analysis

To examine which executive functions accounted for the relation between stereotype threat and math accuracy and stereotype threat and number correct, respectively, we used the same bias-corrected bootstrapping procedure as in Experiment 1.[7] For math accuracy, the direct relation of stereotype threat still predicted math accuracy when performance on the executive functions tasks was taken into account, but it was substantially reduced. Although stereotype threat was related to both the inhibition (Stroop) task and the updating (keep track) task (see above), only the bias-corrected 95% CI for the indirect effect of the updating task on the relation between stereotype threat and math accuracy did not include 0 (−.058 to −.011). The bias-corrected 95% CIs for the indirect effect of the inhibition and shifting tasks included 0 (−.007 to +.002 and −.016 to +.004, respectively; see Figure 2). The model for number correct showed that stereotype threat predicted the number correct when performance on the executive function task was taken into account. However, the bias-corrected 95% CI for the indirect effect of the updating task on the relation between stereotype threat and number correct barely included 0 (−.24 to +.01), and the bias-corrected 95% CIs for the indirect effect of the inhibition and shifting tasks also included 0 (−.02 to +.19 and −.02 to +.11, respectively). These results indicate that only updating, and not inhibition or shifting, mediated the effect of stereotype threat on women's math accuracy. This same pattern was seen for number correct, but was weaker.

Discussion

In Experiment 2, we again showed that stereotype threat's deleterious effect on women's math performance is predicted by updating. Although stereotype threat reduced women's ability to inhibit prepotent responses, neither inhibition nor shifting accounted for the impact of threat on women's math accuracy. By utilizing different measures of updating and shifting and by using a different math task, we were able to provide evidence that the mediational role of updating in accounting for stereotype threat effects on women's performance was probably not due to aspects of the particular measures used to assess the executive functions or to strong similarities between the skills used in the math task and the measure of updating.

Experiment 3

Experiment 3 examined if the effect of stereotype threat on measures of two different outcomes (math performance and risk preference) could be accounted for by different executive functions. In line with Experiments 1 and 2, we attempted to replicate the finding that updating accounted for reductions in women's math performance when experiencing stereotype threat. We also examined if stereotype threat's negative impact on inhibition could account for stereotype threat leading to riskier preferences.

The impact of stereotype threat on risk taking has been examined in only a few studies (e.g., Carr & Steele, 2010; Inzlicht & Kang, 2010), which have yielded somewhat contradictory findings (addressed in detail in the Discussion of this experiment). This past research is couched within a dual-system theory of decision making (e.g., McClure et al., 2004), which argues that there is a quick, associative system and a resource-dependent, deliberative system that are used to make decisions. Because stereotype threat reduces executive functioning, it interferes with the deliberative system and leads people to rely on their associate system when making decisions. Inzlicht and Kang (2010) found that women under threat preferred a riskier lottery than women in a control condition. We used Inzlicht and Kang's measure of risk taking because it presents a choice between a lottery with a low payout but a high probability of winning (Lottery A) and a lottery with a high payout with a low probability of winning (Lottery B). The expected value of the second lottery is higher; therefore, there is clearly a less risky option (Lottery A) and a risky option (Lottery B). Presumably because the risky lottery has a high payout, and the less risky lottery a low payout, people's associative system is initially attracted to it. However, because people realize that the risky lottery offers a low probability of winning and the less risky option a high probability of winning, they will move away from their initial response (i.e., wanting to win big) to a more reasonable one (i.e., wanting to win) using the deliberate system. This decision process seems to map most closely onto the executive function of inhibition, where people must suppress a prepotent response. Thus, we predicted that inhibition would account for the effect of stereotype threat on Inzlicht and Kang's lottery task.

Because we measured inhibition using the Stroop task in Experiments 1 and 2, we utilized a different measure of inhibition in Experiment 3. As noted earlier, this allows us to better control for idiosyncratic effects of a particular measure of executive functions across experiments. Importantly, showing within the same experiment that different executive functions account for different outcomes in response to stereotype threat provides more compelling evidence that examining distinct executive functions can add to our understanding of stereotype threat effects.

Method

Participants. Eighty-two female undergraduates participated for course credit or extra credit. They were randomly assigned to the control condition or the stereotype threat condition. Due to a computer error, data from one participant were incomplete, leaving a final sample of 81 women.

Procedure. Participants were given the same manipulation of stereotype threat used in Experiment 1. They then completed three tasks, one that measured each of the executive functions. As in the previous experiments, the order of these tasks was randomized for each participant. The same measure of updating used in Experiment 1 (letter-memory task) and the same measure of shifting used in Experiment 2 (color–shape task) were used in Experiment 3. We introduced a new measure of inhibition: the antisaccade task.

The antisaccade task consisted of 112 trials. Each trial started with a fixation cross that randomly appeared on the computer screen for 1500 to 3000 ms. After the fixation cross, a box appeared on the right (on half of the trials) or left side (on half of the trials) of the computer screen for 150 ms. Then, one of three numbers (i.e., 4, 7, 9) appeared on the opposite side of the screen from the box for 175 ms. Each number was presented an equal number of times and was immediately followed by a black and white mask. Participants then guessed which number they had seen by pressing the appropriate response key on the number pad. While no response window was imposed on participants during the task, participants were informed in the task instructions to make their "judgments as quickly and accurately as possible." An inhibition index was calculated by dividing the number of trials answered correctly by the total number of trials, with greater inhibition indexed by a greater proportion of trials answered correctly. That is, to the extent that participants can inhibit their tendency to attend to the box, they should be better at identifying the number presented on the opposite side of the computer screen.

Next, participants completed a risk measure (described as being unrelated to the math task) from Inzlicht and Kang's (2010) study. Participants were asked to imagine making a choice between two lotteries that they could play, each costing US$5 to enter. Objectively, one lottery (Lottery B) was relatively more risky (it "has a US$250 prize" and "a 4% chance of winning") than the other lottery (Lottery A; it "has a US$20 prize" and "a 70% chance of winning"). To measure risk, participants indicated their relative preference for the lotteries on a scale ranging from 1 (*Lottery A*) to 9 (*Lottery B*). Participants also completed the same GRE-based math test used in Experiment 2. The order of the risk measure and the math test was randomized. As a manipulation check, participants then completed a three-item measure of threat-based concern (Marx, 2012; $\alpha = .81$; for example, "I worry that if I perform poorly on this test, the experimenter will attribute my poor performance to my gender"), with the responses, ranging from 1 (*strongly disagree*) to 7 (*strongly agree*), on the three items averaged together and greater scores indicating more threat-based concern.

Results

Manipulation check. The manipulation of stereotype threat was effective. Women in the stereotype threat condition showed greater threat-based concern ($M = 3.34$) than women in the control condition ($M = 1.97$), $t(79) = 4.56$, $p < .001$, $d = 1.01$.

Correlations between the executive functions. The inhibition measure was unrelated to the updating measure, $r = .14$, $p = .22$, and the shifting measure, $r = -.06$, $p = .61$. The updating measure was not correlated to the shifting measure, $r = .14$, $p = .20$.

Executive function tasks. The effect of stereotype threat on each of the executive function tasks was examined separately. There was a significant effect of stereotype threat on the inhibition task, $t(79) = -2.34$, $p = .02$, $d = -.50$. Women under threat showed poorer inhibition ($M = 0.52$) than women under the control condition ($M = 0.59$). The results for the updating task also showed an effect of stereotype threat, $t(79) = -2.29$, $p = .03$, $d = -.50$. Women showed poorer updating in the stereotype threat condition ($M = 0.83$) than in the control condition ($M = 0.90$). The stereotype threat manipulation did not affect performance on the shifting task, $t(79) = -.87$, $p = .39$, $d = -.20$ ($M_{\text{Control}} = -203$ ms, $M_{\text{Stereotype threat}} = -147$ ms).

Math performance. Women in the stereotype threat condition showed poorer math accuracy ($M = 0.38$) than those in the control condition ($M = 0.55$), $t(79) = -3.28$, $p = .01$, $d = -.74$. Women under threat also showed a lower number correct ($M = 3.73$) than did those under the control condition ($M = 4.76$), $t(79) = -2.14$, $p = .035$, $d = -.48$. Women in the stereotype threat and control conditions attempted an equivalent number of math problems ($M = 10.28$ and $M = 9.85$, respectively), $t(79) = .50$, $p = .62$, $d = .11$.

Risk taking. Women in the stereotype threat condition rated the riskier lottery as relatively more preferable ($M = 3.70$) than those in the control condition ($M = 2.64$), $t(79) = 2.54$, $p = .01$, $d = .57$. Women under threat also provided their preference more quickly ($M = 6.03$ s) than women not experiencing threat ($M = 6.98$ s), $t(79) = 2.20$, $p = .03$, $d = .49$. Risk taking and math performance were not correlated, $r = -.01$, $p = .90$.

Mediational Analyses

Math performance. To examine which executive functions accounted for the relation between stereotype threat and math accuracy and stereotype threat and number correct, respectively, we used the same bias-corrected bootstrapping procedure as in Experiments 1 and 2. For math accuracy, the direct relation of stereotype threat still predicted the math accuracy when performance on the executive functioning tasks was taken into account, but it was reduced. Although stereotype threat was related to both the inhibition (antisaccade) task and the updating (letter-memory) task (see above), only the bias-corrected 95% CI for the indirect effect of the updating task on the relation between stereotype threat and math accuracy did not include 0 (−.052 to −.003). The bias-corrected 95% CIs for the indirect effect of the inhibition and shifting tasks included 0 (−.015 to +.014 and −.018 to +.002, respectively; see Figure 3). The model for number correct showed that stereotype threat predicted the number correct when performance on the executive function task was taken into account. However, the bias-corrected 95% CI for the indirect effect of the updating task on the relation between stereotype threat and number correct barely included 0 (−.37 to +.02), and the bias-corrected 95% CIs for the indirect effect of the inhibition and shifting tasks also included 0 (−.05 to +.23 and −.08 to +.08, respectively). These results indicate that only updating, and not inhibition or shifting, mediated the effect of stereotype threat on women's math accuracy. As in Experiment 2, this same pattern was seen for number correct, but was weaker.

Risk taking. To examine which executive functions accounted for the relation between stereotype threat and risky choice, we used again a bias-corrected bootstrapping procedure (see Figure 4). The direct relation of stereotype threat no longer predicted risk taking when performance on the executive functions tasks was taken into account. Most importantly, only the 95% bias-corrected CI for the inhibition measure did not include zero (+.011 to +.325), whereas zero was included for both the updating measure (−.153 to +.125) and shifting measure (−.032 to +.147). This analysis shows a significant indirect path whereby stereotype threat leads to

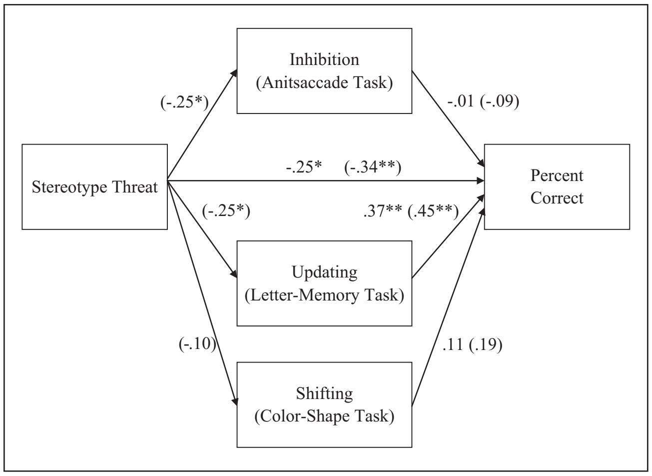

Figure 3. A path diagram showing the effect of stereotype threat on percent of math problems correct as mediated through tasks assessing the executive functions of inhibition, updating, and switching in Experiment 3.
Note. All scores are standardized beta weights. The direct effects between the variables are in parentheses.
$^{*}p < .05$. $^{**}p < .01$.

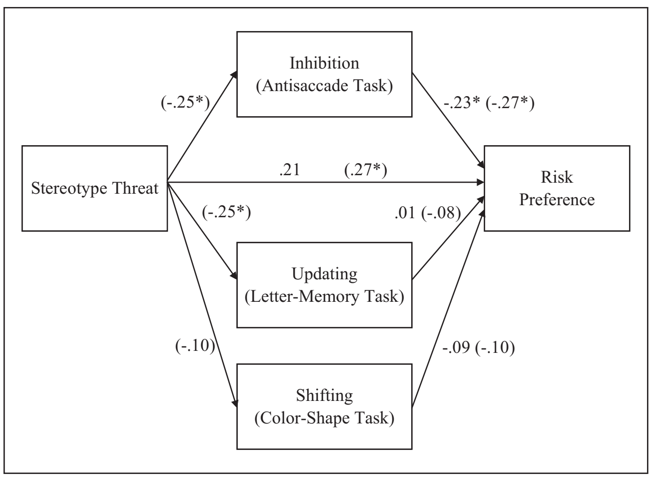

Figure 4. A path diagram showing the effect of stereotype threat on risk preference as mediated through tasks assessing the executive functions of inhibition, updating, and switching in Experiment 3.
Note. All scores are standardized beta weights. The direct effects between the variables are presented in parentheses.
$^{*}p < .05$.

changes in inhibition ability that account for the preference for a risky choice.

Discussion

Experiment 3 replicated the results from the previous experiments by showing that stereotype threat reduced women's math accuracy by impairing updating. However, the effect of stereotype threat on increased risk taking was accounted for by threat's detrimental influence on inhibition. This work highlights that, even within the same experiment, stereotype threat's effect on distinct outcome measures can be uniquely accounted for by different executive functions. This experiment also demonstrates that the influence of stereotype threat on different outcomes can be due to the failures of distinct executive functions.

However, as noted earlier, the previous results regarding the effect of stereotype threat on risk taking are inconsistent. Although in this experiment, and also by Inzlicht and Kang (2010), it was found that stereotype threat increases riskiness, Carr and Steele (2010) found that threat induces risk aversion. We think that several factors might be leading to these seemingly contradictory findings. The lotteries used in these two sets of experiments differ in their expected values; that is, the lottery choices in our work, and that in Inzlicht and Kang's, had different expected values, whereas the choices in Carr and Steele's work had the same expected values. Thus, there was an objectively less risky option in our work, but not in the task used by Carr and Steele, that should be chosen if inhibition is successful. Second, because executive functions are compromised by threat, participants' choices may be biased by associative processes. This processing could lead to different choices based on which aspect of the decision-making task is most salient. In our experiment, there was a large difference in the possible payouts for each lottery option and the possibility of a relatively large reward. Therefore, participants may have been particularly focused on the payout in their decision and, given insufficient deliberative processing, chose the lottery with the "riskier" option. In contrast, Carr and Steele had relatively small payouts (with equal expected values), which may have led participants to focus more on the *probability* of gaining a reward and therefore to choose the option that gave them the greatest chance of gaining a reward, leading to risk aversion. Nonetheless, risk aversion is a normatively irrational behavior, just as choosing the option with the lower expected value is. Therefore, it could be argued that although in terms of "risky" choices these findings seem contradictory, they both demonstrate irrational decision making.

General Discussion

The present work replicated previous stereotype threat research (e.g., Beilock et al., 2007; Inzlicht & Kang, 2010) and provided further evidence for the general theory that reduced executive functioning mediates the relationship between stereotype threat and reduced performance or increased risk taking. Importantly, this work begins to disentangle exactly which of the executive functions are impaired by stereotype threat and the implications of that impairment. Although the executive functions of both inhibition and updating were negatively impacted by stereotype threat, only updating was found to account for women's threat-related decrements in math performance. However, only reduced inhibition as a result of threat accounted for increased risk taking, demonstrating that not all stereotype threat effects are mediated by updating.

These findings highlight how different aspects of executive functioning can account for different outcomes due to stereotype threat and speak to the benefits of utilizing research and theory on executive functioning to hone our predictions, as this work postulates how different executive functions can be delineated and why they are important for understanding cognitive functioning (e.g., Baddeley & Logie, 1999; Miyake et al., 2000). These more nuanced models of executive functioning can help to better explain the various pathways through which threat affects responses on different types of tasks. Elucidating these pathways will provide a richer understanding of the mechanisms underlying threat-based effects and the ability to predict how and why threat can impact a wider range of outcomes.

Our approach could predict which executive functions will account for outcome differences due to the particular aspects of the experimental tasks. Considering that stereotype threat can have a deleterious effect in many different performance domains and on a variety of types of judgments and behaviors (e.g., Inzlicht et al., 2006; Inzlicht & Kang, 2010), some tasks may rely disproportionately on one executive function or some combination of executive functions. This reasoning was tested in Experiment 3, finding that math performance under threat was more dependent on updating, whereas risk taking while under threat was accounted for by inhibition; this reasoning can also be applied to other outcomes examined in the stereotype threat literature. For example, increased aggression following stereotype threat may be due to attempts to control or suppress prepotent urges to aggress (Inzlicht & Kang, 2010). Inhibition may be the executive function most likely to be implicated as a mechanism in this particular effect of stereotype threat. Clarifying the contributions of different executive functions in the relationship between stereotype threat and impaired performance is extremely important as it allows for more specific predictions about the mechanisms underlying stereotype threat effects to be made a priori.

In addition to the theoretical contributions this work could make to the literature pertaining to executive functioning's role in stereotype threat effects, our findings have practical implications for any situations in which stereotyped group members need to perform. For instance, the results for math performance were found across not only a novel math task (MA), but also a task made up of more commonly used math problems (GRE). At the very least, our results suggest that updating is important in many of the types of problems commonly used to assess math performance in education and real-world evaluation situations. In addition, to the extent that the majority of math tasks used to assess performance in these real-world situations rely on updating, women may be at a particular disadvantage due to stereotype threat's effect on updating. It is possible though that a different type of math task that does not rely on updating may provide women with a better opportunity to showcase their ability in these performance settings.

Furthermore, reduced inhibition in response to stereotype threat accounted for increased risk taking; this finding indicates that stereotype threat may lead to other types of risky behaviors such as risky sexual behaviors and drug and alcohol use. Stereotype threat, to the extent it reduces inhibition, may lead people to act more impulsively (e.g., Hofmann et al., 2008) across many different domains (e.g., overeating), failing to resist temptations. Inhibition has also been implicated as a barrier to achieving long-term goals, such as academic achievement (e.g., Eigsti et al., 2006), and, therefore, stereotype threat may, if present in many situations, impact these important outcomes.

This work also suggests a new framework within which interventions that aim to mitigate stereotype threat can be evaluated and applied. Specifically, testing how manipulations that have been previously shown to reduce threat (e.g., self-affirmation, "gender fair" information, learning about threat) have their influence on specific executive functions could aid in our theoretical understanding of how these manipulations work and provide practical suggestions as to under which conditions each way of mitigating threat may be most successful. Furthermore, an extension of our work could consider whether the training of specific executive functions differentially buffers women from different outcomes of stereotype threat. It is possible that increasing updating ability could eliminate math performance decrements, and increasing inhibition could reduce risk-taking behavior. This future work should also examine if executive function training is sufficient to reduce the negative consequences of stereotype threat or if the threat itself must be eliminated from the situation.

Because of stereotype threat's widespread impact, refining our understanding of the various processes by which stereotype threat affects many different outcomes is helpful in allowing researchers to pinpoint and make specific predictions regarding the mechanisms responsible for and types of tasks most impacted by stereotype threat. Future research will benefit from this more detailed way of thinking about how stereotype threat harms different executive functions and how deficits in these executive functions translate into behavior.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

Notes

1. And, some stereotype threat effects are not even mediated by executive functioning (Schmader & Beilock, 2012).
2. In all experiments, the control or the stereotype threat instructions (consistent with random assignment) were reiterated before the instructions of each executive function task. Furthermore, the response accuracy on executive function measures where response latency was used for scoring was greater than 92%.
3. It is notable that the zero-order correlations between measures of executive functions in our experiments are very similar to those obtained by Miyake et al. (2000, Appendix A). Thus, they do not, as they did not in that work, indicate a lack of unity in executive functioning.
4. For this analysis, gender (women = 0, men = 1) and stereotype threat (control = 0, stereotype threat = 1) were recoded and their interaction (multiplicative function) was calculated. The effects of stereotype threat and gender were covariates in this model. The mediational results in all of the mediational analyses presented in this article used 1,000 bootstrap resamples. We ran all of the analyses with 10,000 bootstrap resamples and the additional resamples did not alter the results. We also ran the mediational models without using shifting as a mediator, and the results were unchanged.
5. Multicollinearity was not problematic for any of the mediational models presented in these experiments. The regression equations showed the variance inflation factors below 1.17 and tolerance scores above .857.
6. Participants were also not able to go back to problems that they had previously seen. These measures gave us a good sense, especially when taken along with the amount of time to press the "skip" button and the usage of the scratch paper, of how many problems were attempted. We went back through and examined these data to make sure that the number attempted accurately reflects participants working on a problem—the skip button was used correctly by our participants: when they skipped, they almost always, >90%, skipped problems within 6 s.
7. For the analyses in Experiments 2 and 3, stereotype threat was recoded: control = 0 and stereotype threat = 1.

References

Baddeley, A. D., & Logie, R. H. (1999). Working memory: The multiple component model. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 28-61). New York, NY: Cambridge University Press.

Baumeister, R. F. (1998). The self. In D. Gilbert & S. T. Fiske (Eds.), *The handbook of social psychology* (4th ed., Vol. 1, pp. 680-740). New York, NY: McGraw-Hill.

Baumeister, R. F., & Vohs, K. D. (Eds.). (2004). *Handbook of self-regulation: Research, theory, and applications*. New York, NY: Guilford Press.

Beilock, S. L., Rydell, R. J., & McConnell, A. R. (2007). Stereotype threat and working memory: Mechanisms, alleviation, and spillover. *Journal of Experimental Psychology: General*, *136*, 256-276.

Carr, P., & Steele, C. (2009). Stereotype threat and inflexible perseverance in problem solving. *Journal of Experimental Social Psychology*, *45*, 853-859.

Carr, P., & Steele, C. (2010). Stereotype threat affects financial decision making. *Psychological Science*, *21*, 1411-1416.

Conway, A. R. A., Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilhelm, O., & Engle, R. W. (2005). Working memory span tasks: A methodological review and user's guide. *Psychonomic Bulletin & Review*, *12*, 769-786.

Eigsti, I.-M., Zayas, V., Mischel, W., Shoda, Y., Ayduk, O., Dadlani, M. B., . . .Casey, B. J. (2006). Predicting cognitive control from preschool to late adolescence and young adulthood. *Psychological Science*, *17*, 478-484.

Friedman, N. P., Miyake, A., Corley, R. P., Young, S. E., DeFries, J. C., & Hewitt, J. K. (2006). Not all executive functions are related to intelligence. *Psychological Science*, *17*, 172-179.

Hofmann, W., Gschwendner, T., Wiers, R., Friese, M., & Schmitt, M. (2008). Working memory capacity and self-regulation: Towards an individual differences perspective on behavior determination by automatic versus controlled processes. *Journal of Personality and Social Psychology*, *95*, 962-977.

Hofmann, W., Schmeichel, B. J., & Baddeley, A. D. (2012). Executive functions and self-regulation. *Trends in Cognitive Sciences*, *16*, 174-180.

Inzlicht, M., & Kang, S. K. (2010). Stereotype threat spillover: How coping with threats to social identity affects, aggression, eating, decision-making, and attention. *Journal of Personality and Social Psychology*, *99*, 467-481.

Inzlicht, M., McKay, L., & Aronson, J. (2006). Stigma as ego-depletion: How being the target of prejudice affects self-control. *Psychological Science*, *17*, 262-269.

Jamieson, J. P., & Harkins, S. G. (2007). Mere effort and stereotype threat performance effects. *Journal of Personality and Social Psychology*, *93*, 544-564.

Marx, D. M. (2012). Differentiating theories: A comparison of stereotype threat and stereotype priming effects. In M. Inzlicht & T. Schmader (Eds.), *Stereotype threat: Theory, process, and application* (pp. 124-140). New York, NY: Oxford University Press.

Masicampo, E. J., & Baumeister, R. F. (2008). Toward a physiology of dual-process reasoning and judgment: Lemonade, willpower, and expensive rule-based analysis. *Psychological Science*, *19*, 255-260.

McClure, S. M., Laibson, D. I., Lowenstein, G., & Cohen, J. D. (2004). Separate neural systems value immediate and delayed monetary rewards. *Science*, *306*, 503-507.

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive Psychology*, *41*, 49-100.

Morris, N., & Jones, D. M. (1990). Memory updating in working memory: The role of the central executive. *British Journal of Psychology*, *81*, 111-121.

Nosek, B. A., Banaji, M. R., & Greenwald, A. G. (2002). Math=male, me=female, therefore math not ≠ me. *Journal of Personality and Social Psychology*, *83*, 44-59.

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

Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior Research Methods*, *40*, 879-891.

Rydell, R. J., McConnell, A. R., & Beilock, S. L. (2009). Multiple social identities and stereotype threat: Imbalance, accessibility, and working memory. *Journal of Personality and Social Psychology*, *96*, 949-966.

Schmader, T., & Beilock, S. L. (2012). Mechanisms: An integration of processes that underlie stereotype threat. In T. Schmader & M. Inzlicht (Eds.), *Stereotype threat: Theory, process, and application* (pp. 34-50). New York, NY: Oxford University Press.

Schmader, T., & Johns, M. (2003). Converging evidence that stereotype threat reduces working memory capacity. *Journal of Personality and Social Psychology*, *85*, 440-452.

Schmader, T., Johns, M., & Forbes, C. (2008). An integrated process model of stereotype threat on performance. *Psychological Review*, *115*, 336-356.

Schmiedek, F., Hildebrandt, A., Lövdén, M., Wilhelm, O., & Lindenberg, U. (2009). Complex span versus updating tasks of working memory: The gap is not that deep. *Journal of Experiment Psychology: Learning, Memory, and Cognition*, *35*, 1089-1096.

Steele, C. M., Spencer, S. J., & Aronson, J. (2002). Contending with group image: The psychology of stereotype and social identity threat. In M. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 34, pp. 379-440). New York, NY: Academic Press.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, *18*, 643-662.