STEREOTYPE SUPPRESSION IN HIGH VERSUS LOW STIGMA-CONSCIOUS WOMEN EXPERIENCING STEREOTYPE THREAT

Sunyoung Hwang

Thesis Submitted in Partial Fulfillment of the Requirements

for the Bachelor of Arts Degree at Hamilton College

May 13, 2013

Advisor: Jennifer Borton

Abstract

Activating the stereotype of women's lesser ability in math during test taking causes women to underperform on math tests, a phenomenon known as stereotype threat. The present study examined how women high (HSC) versus low (LSC) in stigma consciousness differ in their experience of stereotype threat and their use of stereotype suppression. Given that stereotype threat effects are exacerbated for HSC women, HSC stereotype-threatened women were expected to underperform on a given math-task relative to their LSC counterparts. They were also expected to suppress the female stereotype to a greater extent prior to the task, and to suffer from greater post-suppressional rebound afterwards. Following a stereotype threat manipulation, in which only those in the threat condition were told that an upcoming task was indicative of math ability, participants completed a working memory test. As expected, HSC stereotype-threatened participants demonstrated test underperformance; however, they neither suppressed the female stereotype nor experienced post-suppressional rebound. The findings strongly suggest that HSC women are, in fact, more vulnerable to stereotype threat than are LSC women, who in turn appear to be protected against stereotype threat even when faced with explicit cues of prejudice.

Stereotype Suppression in High versus Low Stigma-Conscious Women Experiencing Stereotype Threat

Despite making up almost half of the overall American workforce, women make up less than a quarter of the science, technology, engineering, and math (STEM) workforce (Beede, Julian, Langdon, McKittrick, Khan, & Doms, 2011). This underrepresentation has serious negative implications in our current job market, in which the high-growth sectors are becoming increasingly math and technology-oriented (Dickler, Lee, & Swiatek, 2011) and STEM jobs are among the most profitable (Langdon, McKittrick, Beede, Khan, & Doms, 2011) and gender-fair in terms of wage (Beede et al., 2011).

Since STEM workers typically attain high-level degrees (Langdon et al., 2011), it is unsurprising that the gender gap found at the professional level is mirrored in postsecondary education. At the doctorate level, women are grossly underrepresented in mathematics and the physical sciences, with less than one-third of the doctorate degrees in these fields being awarded to women. Women make up an even smaller percentage of earned doctorates in engineering and computer science (National Science Fund, 2011).

Shedding light on women's underrepresentation in these quantitative fields is their underachievement at high-level math. A recent meta-analysis by Lindberg, Hyde, Petersen, and Linn (2010) suggests that the gender performance differential in pre-college math has dissipated over the past two decades, dispelling the stereotype that men are superior to women at math. However, despite being on par with men, female students are still underrepresented in university high-level math courses. Moreover, belief in the negative math-gender stereotype still thrives among students in these courses (Good,

Aronson, & Harder, 2008). Given that these courses act as gateways to careers in the quantitative fields (Good et al., 2008), women's lack of support and achievement in these high-level math courses bar their further participation in mathematics and related areas.

In order to make sense of the math-gender disparity within and beyond academia, some theorists have looked to broad sociocultural forces, such as discrimination, to explain why women have not yet achieved their potential in STEM fields (e.g., Ceci & Williams, 2009). On the contrary, other theorists, namely stereotype threat theorists, have looked to the pervasive math-gender stereotype as an explanation for women's math underperformance (Steele & Aronson, 1995). Unlike broad sociocultural approaches, stereotype threat theory offers a strictly situational account for why women underperform in the math domain. The present study examines a potential individual differences factor that may impact a woman's experience and response to stereotype threat.

Stereotype Threat

All invested math students feel pressure to perform well on math tests, but women bear an additional pressure that men do not. Unlike men, women must perform well in the face of the existing negative stereotype about their gender's math ability. If a female student's math performance is substandard, she runs the risk of confirming or being judged by the relevant math-gender stereotype. The added burden of having to contend with this stereotype may explain why women underperform at high-level math. This dilemma is known as stereotype threat (Steele & Aronson, 1995).

Stereotype threat describes an immediate situational threat, or "threat in the air" (Steele, 1997), that is perceived when a test taker is aware that his or her performance may be evaluated with respect to the stereotype. According to stereotype threat theory,

perceiving this threat compromises the test taker's ability, thereby causing underperformance. Spencer, Steele, and Quinn (1999) first demonstrated the negative effects of stereotype threat on women's math performance using what has become the classic stereotype threat manipulation. By simply informing participants that a difficult math test demonstrates gender differences in performance, women in this "threat" condition underperformed relative to men, whereas women informed that the test was gender-fair performed no differently than men. The effectiveness of manipulating test presentation to induce stereotype threat speaks to the situational nature of stereotype threat; it arises from the performance situation itself.

Given its situational nature, it follows that stereotype threat is not specific to a particular social group or task domain. A test taker is capable of experiencing stereotype threat insofar as he or she belongs to a group that is socially expected to underperform (Steele, 1997). The broad scope of stereotype threat has been supported by its numerous demonstrations in a variety of social groups and task domains, including Blacks (Steele & Aronson, 1995), Latinos (Gonzales, Blanton, & Williams, 2002; Schmader & Johns, 2003), and those of low socioeconomic status (Croizet & Claire, 1998) on intellectual tasks; Whites on tests of natural athletic ability (Stone, Lynch, Sjomeling, & Darley, 1999; Stone, 2002); and, of particular interest to the present paper, women on math tasks (Brown & Pinel, 2003; Carr & Steele, 2009; Davies, Spencer, Quinn, & Gerhardstein, 2002; Good, Aronson, & Harder, 2008; Kiefer & Sekaquaptewa, 2007; Logel, Iserman, Davies, Johns, Schamder & Martens, 2005; Quinn, & Spencer, 2009; Spencer, Steele & Quinn, 1999). Stereotype threat has even been demonstrated in social groups that are not traditionally stigmatized, such as White men. In a study by Aronson, Lustina, Good,

Keough, Steele, and Brown (1999), informing White male students that their math test performance would be compared to that of Asian male students, who are stereotypically superior at math, successfully provoked stereotype threat effects. Again, this finding reinforces the fact that, regardless of their past experience in the test domain, test takers can feel threatened by a stereotype when it is made relevant during the testing situation.

Targets of stereotype threat. However, not all people who encounter stereotype-threatening testing conditions experience stereotype threat to the same extent. A person must satisfy several conditions in order to be susceptible to its negative effects. The most basic of these conditions is stereotype awareness. That is, people must be aware of the relevant stereotype in order to perceive the risks associated with stereotype-consistent performance. Although people must be aware of the stereotype, they do not need to subscribe to the stereotype, nor personally satisfy the stereotype in order to be threatened by it (Steele, 1997); people can still fear confirming a stereotype they know is untrue of themselves or in general. What is necessary, however, is an understanding that others in the testing situation may assess their performance according to the relevant stereotype.

Additionally, people must identify with the stereotyped task domain in order to feel threatened by the task-relevant stereotype. Domain identification describes the perceived importance of a task domain to self-definition (Aronson et al., 1999). Since those who strongly identify with a task domain perceive task ability as important to their self-concept, it follows that a part of their self-regard is contingent on task performance. Therefore, domain identification is a necessary condition because it disposes people to the performance pressure characteristic of stereotype threat. The importance of domain identification has been supported by studies in which those who highly achieve within the

stereotyped task domain exhibit stereotype threat effects. Despite clearly refuting the relevant stereotype, high-achievers are not only capable of experiencing stereotype threat (Aronson et al., 1999; Good et al., 2007) but may be especially susceptible to its effects (Steele, 1997). This vulnerability may be due to the fact that high achievement within a task domain involves high identification with that domain.

Although gender identification has also been recognized as a moderator of math-gender stereotype threat (Schmader, 2002), such that only women who sufficiently value being a woman are threatened by the female stereotype, this may depend on how the threat is framed. In a study by Wout, Danso, Jackson, and Spencer (2008), gender identification only moderated stereotype threat effects in a group-threat condition, in which test takers were told that their math test performance would inform the researchers about gender differences in math performance. Conversely, women in a self-threat condition, in which they risked confirming the relevant stereotype to only themselves, underperformed irrespective of their gender identification. Therefore, the extent of a woman's gender identification may play a role only when the math-gender threat is oriented towards women as a whole, and not them individually (Brown & Pinel, 2003; Kiefer & Sekaquaptewa, 2007; Wout et al., 2008).

Stereotype threat tasks. Just as there are necessary conditions for a person to experience stereotype threat, so are there necessary conditions for a task to be sensitive to its detrimental effects. In order for a test to demonstrate performance decrements, the stereotype must be applicable to the test itself (Wheeler & Petty, 2001). The necessity of stereotype applicability is supported by findings that math-gender stereotype threat

impedes performance only on math tests and not on verbal tests, for which women are not stereotyped (Logel, Walton, Spencer, Iserman, & Von Hippel, 2009).

Furthermore, the test must be described as or assumed to be diagnostic of the stereotyped ability. Presenting a task as diagnostic induces stereotype threat because it creates the necessary evaluative context in which stereotypic assessment is possible. Diagnosticity of a task can also be implied by its difficulty. Stereotype threat is more likely to occur under difficult testing conditions because tests that require one's highest effort are assumed to be diagnostic and therefore subject to stereotypic evaluation (Logel et al., 2009; Spencer et al., 1999). But, as demonstrated by the classic stereotype threat manipulation, these diagnostic conditions are threatening only if the relevant negative stereotype is made salient. Past research has shown that the manner of stereotype activation can vary widely, from the blatant to the subtle (Stone & McWheenie, 2008). Relatively blatant stereotype activation has included directly informing test takers that their social group underperforms on this test (e.g., Aronson et al., 1999) or implying it by stating that the test produces group performance differences (e.g., Brown & Pinel, 2003). More subtle stereotype activation operates through making the participant's social group identity salient, which has been done in a number of ways, including completing a demographics form prior to testing (Steele & Aronson, 1995), being the only in-group member present (Inzlicht & Ben-Zeev, 2000; Sekaqueptewa & Thompson, 2003), being exposed to stereotypic group portrayals (Davies et al., 2002), and interacting with an out-group member (Danso & Esses, 2001; Logel et al., 2009).

Underlying processes. Negative performance following stereotype threat has been partially explained by several processes, all of which are initiated by the activation

of stereotypic constructs. It is important, however, to distinguish stereotype threat from another model that touches on stereotype activation: the ideomotor model.

The ideomotor phenomenon describes the automatic tendency for behavior to follow suit with perceptual representations (Wheeler & Petty, 2001). That is, simply perceiving or thinking about an action or concept can prime a person to behave correspondingly. For example, people may unconsciously imitate the facial expressions or vocal patterns of a person they have recently seen (Dijksterhuis & Bargh, 2001). When applied to stereotype activation, the ideomotor model states that activating stereotypic constructs will result in stereotype-consistent behavior. Since these priming effects can occur without conscious awareness, stereotype threat and ideomotor theory differ in a critical way: the ideomotor mechanism does not rely on the test taker's conscious feelings or motivational states to explain underperformance (Wheeler & Petty, 2001). Although this priming theory can explain phenomena caused by the activation of an *other*-stereotype, such as why activating the supermodel stereotype causes people to perform poorly on intellectual tasks (Dijksterhuis & van Knippenberg, 1998), or why activating the stereotype of the elderly causes people to walk more slowly (Bargh, Chen, & Burrows, 1996), it cannot explain why activating a negative *self*-relevant stereotype results in task underperformance.

Davies, Spencer, Quinn, and Gerhardstein (2002) distinguished stereotype threat effects from ideomotor effects in a study in which men and women were exposed to commercials depicting the female stereotype prior to taking a math test. Although the female stereotype was activated in both men and women, as indicated by results of an Implicit Association Test (IAT), only women underperformed on the math test. Since the

ideomotor model purports that the activation of any stereotype, regardless of self-relevance, results in stereotype-consistent behavior, this model cannot explain women's underperformance. Instead, women's emotional and cognitive responses should be taken into account, as they are under the stereotype threat model.

According to the theory of stereotype threat, it is not stereotype activation in itself that causes underperformance, but the affective and cognitive responses to this stereotype activation. Affective responses, such as increased arousal (Ben-Zeev, Fein, & Inzlicht, 2005; O'Brien & Crandall, 2003) and heart rate (Croizet, Despres, Guizins, Huguet, Leyesns, & Meot, 2004), are collectively understood as a physiological stress response (Schmader, Johns, & Forbes, 2008). This stress response has been found to strain cognitive efficiency (Croizet et al., 2004). Further straining the test taker's ability are the concerns and worries that accompany the stress response. Ruminating on the negative thoughts associated with the stereotype can co-opt necessary cognitive resources, namely working memory capacity (Beilock & Beilock, 2007). By absorbing a considerable amount of cognitive resources, either physiologically or cognitively, the negative thoughts activated during stereotype threat (Schmader & Johns, 2003; Schmader, Johns, & Forbes, 2008) leave behind only a limited amount of cognitive resources with which to complete the actual task. Thus, stereotype threat theory attributes task underperformance to the burdened cognitive capacity of the stereotype threatened.

Long-term consequences of stereotype threat. Task underperformance is not the only consequence of stereotype threat. In the short run, stigmatized people who receive negative feedback on a stereotyped task may situationally "disengage" from the stereotyped task domain (Nussbaum & Steele, 2007), during which their self-esteem is

momentarily suspended from the performance outcomes of the stereotyped task (Major, Spencer, Schmader, Wolfe, & Crocker, 1998). Since situationally disengaged people can insulate themselves from negative feedback while still identifying with the task domain, this temporary disengagement may be an adaptive coping strategy (Nussbaum & Steele, 2007). However, repeated experiences of stereotype threat may eventually lead stigmatized people to disengage to the point that the domain is no longer a significant part of their self-concept or, in other words, to disidentify from the domain (Osborne, 1997; Steele, 1997). This psychological withdrawal from the stigmatized domain leads to underparticipation (Steele, 1997) as well as poor performance and motivation (Pinel, Warner, & Chua) within that domain. But certain stigmatized people are more predisposed to experience stereotype threat, and are therefore more at risk for domain disidentification, than are others (Brown & Lee, 2005; Pinel et al., 2005). One example of such an individual differences factor is stigma consciousness.

Stigma Consciousness

People differ in their experiences of stigmatization. Whereas some people do not allow a self-stigma to affect their experiences, others are constantly aware of, and therefore affected by, this self-stigma. The chronic awareness of one's stigmatized status is known as stigma consciousness (Pinel, 1999). Stigma consciousness is not simply being aware of a self-stigma, as it is possible for a person to be aware of one's stigmatized status and not be affected by it. Rather, stigma consciousness describes a fixation on one's stigmatized status.

By being self-conscious of their stigmatized status, people high in stigma consciousness interpret interactions and experiences from a perspective that emphasizes

their self-stigma (Pinel, 1999). As such, highly stigma conscious (HSC) people are constantly anticipating, and are therefore vigilant for cues of prejudice. These chronic prejudice expectations affect how people interpret interactions with out-group members. For example, relative to low stigma conscious (LSC) women, HSC women are more likely to anticipate negative interactions with men (Pinel, 2004) and, when actually faced with these negative interactions, are more likely to attribute them to discrimination (Pinel, 1999; Pinel, 2002).

But the heightened vigilance of HSC women for prejudice may also operate on an unconscious level. In a study by Kaiser, Vick, and Major (2006), women's chronic expectations for prejudice predicted their preconscious attention to sexist words during a Stroop task. Furthermore, women's chronic prejudice expectations predicted preconscious attention only when the female stereotype was made salient during the testing situation, as it is during experiences of stereotype threat. Given that HSC women are more vigilant than LSC women for stereotype cues in threatening environments, it follows that they are more vulnerable to stereotype threat. Brown and Pinel (2003) demonstrated this vulnerability by informing both HSC and LSC women that an upcoming math test produced gender differences. However, even after this stereotype threat manipulation, only HSC (but not LSC) women underperformed as a result of stereotype activation.

Brown and Pinel (2003), however, have suggested that using an extreme stereotype threat manipulation, in which the self-relevant stereotype is explicitly mentioned, may successfully induce stereotype threat in LSC women. By explicitly informing women that their group is expected to underperform, even women who are

otherwise unaware of the risks associated with stereotypic performance should feel burdened by the now salient stereotype. While this particular manipulation has not yet been used in stereotype threat research, Pinel (2004) successfully demonstrated that LSC women can be situationally manipulated to behave as HSC women if they are asked to reflect upon past stigmatization. When asked to focus on previous experiences of discrimination on the basis of their gender, women low in dispositional (trait) stigma consciousness adopted a situationally induced (state) stigma consciousness. As a consequence of this temporarily elevated stigma consciousness, women typically low in stigma consciousness adopted the same attributional style as women chronically high in stigma consciousness.

Coping with stereotype threat. As previously mentioned, people disidentify from stigmatized domains if they continuously experience stereotype threat. The interplay between stigma consciousness, stereotype threat, and domain disidentification has been demonstrated within academic settings, as level of stigma consciousness has been found to predict the GPA (Brown & Lee, 2005) as well as the extent of psychological disengagement from the academic domain (Pinel et al., 2005) of academically stigmatized students (e.g., Blacks, Latinos) but not of non-stigmatized students (e.g., Whites, Asians). Thus, as a consequence of their predisposition to stereotype threat, HSC students are at greater risk than LSC students for academic disidentification. While some stereotype-threatened people cope with stereotype threat by psychologically removing themselves from stigmatized domains, others may rely on a more immediate coping strategy, one that occurs during test taking: stereotype suppression.

Stereotype Suppression

Those who attempt stereotype suppression employ it as a means of clearing the mind of negative stereotypic thoughts, as they believe it will allow them to focus on the task at hand. However, due to the nature of the cognitive processes involved, thought suppression leads to unintended negative effects. One negative effect is the reduction of working memory capacity (Schmader & Johns, 2003). This reduction can be attributed to the high mental control required by stereotype suppression.

As described by Wegner and Erber (1992), thought suppression is a taxing procedure involving two simultaneous cognitive processes: a controlled process responsible for self-distraction, and an automatic process responsible for detecting the undesired thought. At the onset of thought suppression, the controlled process engages in a "controlled distractor search," through which thoughts unrelated to the target are sought out as potential distractors. Once an appropriate distractor is found, conscious attention shifts from the search to the selected distractor. But this distraction is only momentary, as it is eventually disrupted by the intrusion of the target. With the reappearance of the target thought, the controlled distractor search is reinitiated. Sustaining this cycle between searching and self-distracting is the automatic process. While the controlled process provides necessary distractors, the automatic process continually detects any signs of an emerging target through an "automatic target search." In contrast to the controlled distracter search, this target search is continuous and beyond any conscious control or awareness.

Paradoxically, the very intrusions that the controlled and automatic processes work to avoid are, in part, due to this automatic target search. That is, in order to detect

the target thought, the target thought must be kept activated, albeit unconsciously. Therefore, despite attempts to purge the mind of the target, thought suppression maintains it in trying to detect it. Moreover, the detection of a budding target thought inevitably draws conscious attention to it, consequently transforming an unconscious trace into a full-fledged intrusion.

These searches, and the constant switching between them, are cognitively demanding. So, although thought suppression appears to free the mind of the burden of a salient stereotype, it limits cognitive capacity in itself (Schmader & Johns, 2003). Since thought suppression is already cognitively taxing, an additional cognitive load, such as that imposed by a time constraint or a concurrent task, undermines the success of thought suppression (Wegner & Erber, 1992). When deprived of necessary resources, the controlled process is incapable of finding distractors and the automatic process goes unchecked (Wegner, 1994). Consequently, the unceasing automatic target search repeatedly detects the target thought, resulting in widespread activation of the initially suppressed thought. Spencer demonstrates this suppression failure in a study in which placing stereotyped-threatened women under an additional cognitive load led to increased activation of stereotype-associated concepts (as cited in Schmader & Johns, 2003, p. 450).

Similarly, ceasing a stereotype suppression attempt causes previously suppressed thoughts to become hyperaccessible (Wegner, Schneider, Carter, & White, 1987). This post-suppressional rebound has been demonstrated with the suppression of nonself-stereotypes as well as self-stereotypes. In a study by Macrae, Bodenhausen, Milne, and Jetten (1994), participants performed an impression formation task in which they were

provided a picture of a Black man and asked to write about a typical day in his life. Consistent with the rebound effect, the stories of participants who had been instructed to suppress the Black stereotype contained greater stereotypical content than did the stories of those who had not been instructed to suppress.

Interestingly, similar results have been found without using explicitly instructing participants to thought suppress. In an experiment by Wyer, Sherman, and Stroessner (1998), participants performed the same impression formation task as that used by Macrae and colleagues (1994), only this time they were not instructed to suppress. Instead, participants were either told that the study was conducted by the organization, "African Americans for Intercultural Understanding," or were given no background at all. Those told about the organization self-initiated stereotype suppression, as this information acted as a cue to rein in prejudiced responses. Consistent with previous research, the stories written by participants who spontaneously suppressed were greater in stereotypicality than were those of control participants.

These situational cues have also prompted spontaneous thought suppression during experiences of stereotype threat. In one study by Logel, Walton, Spencer, and Iserman (2009), female engineering students who interacted with a sexist male colleague prior to an engineering test spontaneously suppressed the female stereotype prior to testing. This suppression was indicated by these women's slower reaction times to stereotypic words on a lexical decision task taken before the test, relative to control participants' reaction times. Logel, Isermen, Davies, Quinn, and Spencer (2009) replicated these results with female math students and, by using a lexical decision task

following the math test, also found evidence for the subsequent post-suppressional rebound effect as well.

If stereotype suppression is a seemingly rational coping strategy, there is reason to expect that both HSC and LSC women will spontaneously suppress the female stereotype while experiencing stereotype threat. Although it is possible that HSC women are more prone to suppressing stereotypical thoughts than are LSC women, this is unlikely the case. Once people are led to experience stereotype threat, the relevant stereotype is activated in these people regardless of their chronic sensitivity to stereotypic constructs. Thus, assuming that the female stereotype is activated in both HSC and LSC stereotype threatened women, both groups should look to thought suppression as a means of improving test ability. However, this neither means that the stereotype will be activated to the same extent in HSC and LSC women, nor that their stereotype suppression will produce the same outcomes. Although there is no direct research that compares the spontaneous stereotype suppression of HSC and LSC women, there is some research that suggests that the female stereotype is more strongly activated in stereotype-threatened HSC women than in their LSC counterparts, and that, as a consequence, they will suffer greater post-suppressional rebound than will LSC women (e.g., Kiefer & Sekaquaptewa).

In a study by Kiefer and Sekaquaptewa (2007), HSC women were found to possess strong gender-math stereotypic associations, as measured by an Implicit Associations Test. Moreover, these associations remained strongly activated during math testing even when the test was described as non-diagnostic, suggesting that the stereotype is chronically accessible for HSC women. In another study, in which both HSC and LSC women were randomly assigned to either suppress or not suppress the stereotype of

women's poorer spatial skills while interacting with a male confederate, HSC women in the suppression condition displayed more stereotypic behavior (e.g., less dominant nonverbal behavior) in their interactions than did their LSC counterparts. These findings suggest that the unsuccessful stereotype suppression of HSC women may be due to their more accessible stereotypic constructs (Borton, Reiner, Vazquez, Ruddiman, & Anglin, 2011). Given that the female stereotype is chronically accessible for HSC women, their stereotype hyperaccessibility following directed suppression is to be expected. However, little is known about whether there are differences between HSC and LSC women in self-initiated stereotype suppression, and whether HSC women will suffer greater costs during suppression due to the strongly activated female stereotype.

The Present Study

The present study examined how women high versus low in stigma consciousness differ in their experience and response to stereotype threat. First, I assessed whether LSC women would also experience stereotype threat when faced with an extreme math-related threat. Given past demonstrations that LSC women can be situationally manipulated to behave as HSC women, I expected that when the math-gender stereotype was explicitly activated during a testing situation, both HSC and LSC women would experience stereotype threat. Thus, I predicted that, regardless of stigma consciousness level, all women in the threat condition would underperform on a working memory capacity test relative to control participants.

Second, I predicted all participants in the threat condition would spontaneously suppress the female stereotype prior to completing the working memory capacity test. This spontaneous stereotype suppression would be indicated by threat participants'

slower reaction times to female stereotypic words relative to control participants on a lexical decision task prior to testing. Given that the female stereotype is more strongly activated for HSC women, I expected that HSC women to require greater suppression of the stereotype than would their LSC counterparts in order to focus on the test. Thus, I predicted that HSC women would avoid stereotypic constructs to a greater extent than would LSC women, and that they would therefore react more slowly than LSC women to stereotypic words. Moreover, I expected HSC women to require more effort to suppress the highly accessible stereotype, as indicated by self-report.

Third, I expected that this additional effort would compromise the test performance of HSC women in the stereotype threat condition. I therefore predicted that these women would underperform on the working memory capacity task relative to their LSC counterparts. Furthermore, I expected this additional effort, when compounded with the cognitive demand of the test itself, to undermine the success of stereotype suppression. Due to their greater cognitive load, I expected HSC women to exhibit greater stereotype accessibility than their LSC counterparts once they ceased stereotype suppression. Therefore, my fourth hypothesis was that HSC women in the threat condition would react more quickly than LSC women in the threat condition to stereotypic words on a lexical decision task following the test. No such difference was expected in the control condition.

Method

Participants

The participants were 60 female Hamilton College students who majored in mathematics (46.67%), biology (18.33%), economics (15%), biochemistry (10%),

chemistry (6.67%), or physics (3.33%). The participants ranged in age from 19 to 23 years (*M*=21.0, *SD*=.87), and were predominantly White (83.33%). Participants were recruited through e-mail and were compensated with either $10 or 1 extra credit point in a psychology course of their choice.

Materials

Math identification items. The math identification items used on the pre-study questionnaire were the same two statements used by Carr and Steele (2009) to assess math identification ("It is important to me that I do well at math" and "I am good at math"). Participants responded to each item along a 7-point Likert-type scale ranging from (strongly disagree) to 7 (strongly agree). Consistent with previous research, participants were required to score above the midpoint on each math identification item to qualify for the experiment (Aronson et al., 1999; Carr & Steele, 2009).

Stigma Consciousness Questionnaire for Women (SCQ-W). The *SCQ-W* (Pinel, 1999) is a 10-item questionnaire that assesses the extent to which women are self-conscious of their stigmatized status (e.g., "Most men have a lot more sexist thoughts than they actually express," "Stereotypes about women have not affected me personally" (reverse scored)). Participants responded to each item along a 7-point Likert-type scale ranging from 1 (strongly disagree) to 7 (strongly agree), and 7 of the items were reverse scored. The mean of the 10 *SCQ-W* items served as a composite measure of stigma consciousness (Cronbach's $\alpha = .81$).

Operation Span Task (OSPAN). The OSPAN (Turner & Engle, 1989) is a complex span task that measures working memory capacity. On Microsoft PowerPoint

software, participants were presented with a mathematical equation (e.g., 36/3 + 6 = 18) paired with a random word (e.g., bus). Each math-word pair was presented for 8 seconds, during which participants had to decide whether the math equation was correct as they rehearsed the word for later recall. These math-word pairs were presented in series that varied in length, ranging from 2 to 6 math-word pairs. At the end of each series, participants were asked to recall all of the words in the series in the order in which they were presented. Participants were given 3 practice sequences before completing 16 test sequences, for a total of 60 correct words. Score on the OSPAN test served as the primary dependent variable.

Lexical decision task (LDT). The lexical decision task was used to assess the activation of the female stereotype by assessing the speed with which participants categorized serially-presented letter strings as either words or nonwords. Participants completed one lexical decision task before and one after the OSPAN; each lexical decision task had different words, and order was counterbalanced across participants. Each task included 10 stereotypic words (e.g., indecisive), 10 neutral words (e.g., wood), and 10 nonwords (e.g., stropline). Stereotypic and neutral words were the same as those used in previous studies (Carr & Steele, 2009; Logel et al., 2009). Words in different conditions were matched for length and frequency in the English language using norms established by Kucera and Francis (1967). Nonwords were created using a technique used by Lacruz and Folk (2004) in which real words were split into "heads" and "bodies" (e.g., stake = st + ake), and the heads and bodies of multiple words not used as either stereotypic or neutral letter strings were randomly recombined to produce nonwords. The words used to create the nonwords were randomly selected from an online word

generator, and nonwords were matched to stereotypic and neutral words in terms of length.

Both lexical decision tasks were administered using PsyScope X software (Cohen, MacWhinney, Flatt, & Provost, 1993). Each letter string was presented in black 80 pt. Helvetica font at the center of a white screen for 1000 ms. Participants indicated whether the letter string was a word or a nonword by pressing either the 'd' key or the 'k' key, respectively. Following a 500 ms delay, the next letter string was signaled by a plus sign ('+') that remained on the screen for 500 ms before it was replaced by the next letter string.

Stereotype suppression. To assess whether participants attempted stereotype suppression, participants answered the following questions: "Are you aware of the stereotype that women are not as good as men at math?", "How important was it for you to do well on the working memory capacity test?", and "While you were taking the test, did you try to avoid thoughts or feelings associated with the gender stereotype about math ability?" (yes/no). If participants reported suppressing the stereotype, they were also asked the following two questions, both assessed on 7-point rating scales: "How hard did you try to avoid thinking about the stereotype?" and, "To what extent did avoiding the stereotype interfere with your test taking ability?".

Procedure

Participants who met the previously described criteria arrived at the lab individually, ostensibly for a study examining how math-oriented people process various types of information. Participants provided informed consent and were randomly assigned to the threat or control condition.

Those in the stereotype threat condition were told that the experimenter was particularly interested in gender differences in processing mathematical information. They were then told that the upcoming working memory capacity test strongly predicted math ability and future success in STEM fields, that men typically outperform women on the test, and that their results would be used to understand why this gender difference exists. In contrast, participants in the control condition were simply reminded that the experimenter was interested in the cognitive processing of math-oriented people. Control participants were then told that test performance was unimportant because the experimenter was only interested in observing what cognitive processes are used during the test, and that the test shows no gender difference in performance; there was no mention of STEM fields or math ability. Participants in both conditions were also told to give their best effort on the working memory capacity test, and that, in addition to the working memory capacity test, they would complete two verbal tasks, one before and one after the test.

Following the stereotype threat manipulation, participants were led to a private room in which they sat in front of a Macintosh OS X version 10.5.8 laptop. All tests and questionnaires were completed under these conditions. Participants began the testing phase with the first lexical decision task, which was described as the first verbal task. Participants were instructed to complete the task as quickly and accurately as possible, and to alert the experimenter when they had finished. Participants then completed the OSPAN. Again, participants were instructed to tell the experimenter when they were done. After they were shown the same instructions as the first lexical decision task, participants completed the second, and final, lexical decision task, thus completing the

testing phase. Participants then completed the stereotype suppression and demographic items. Afterwards, they were thoroughly debriefed, compensated, and thanked for their participation.

Results

Manipulation Checks

On the question assessing whether participants were aware of the math-gender stereotype, 2 out of the 60 participants indicated that they were unaware of the stereotype. Their data as well as the data of 4 other participants who were suspicious of the stereotype threat manipulation were deleted.

Importance of doing well. On the item assessing participants' ratings of the importance of doing well on the OSPAN test, there was a significant main effect for condition, $t(53) = -2.30$, $p = .026$ ($\beta = -.30$), such that participants in the control condition (M=6.04, SD=.79) rated the test as significantly more important than did those in the threat condition (M=5.42, SD=1.21). Neither the main effect of stigma consciousness nor the stigma consciousness x condition interaction was significant.

Stereotype suppression ratings. Again, there was a significant main effect for condition, $t(53) = -2.25$, $p = .029$ ($\beta = .30$), such that those in the threat condition reported suppressing the stereotype to a greater extent during the OSPAN test than did those in the control condition. Neither the main effect of stigma consciousness nor the stigma consciousness x condition interaction was significant. In addition, there were no significant predictors of participants' ratings of how effortful or disruptive stereotype suppression was.

Effects of Stereotype Threat on Working Memory

Prior to entry in regression equations, stigma consciousness was centered around its mean (M=4.16, SD=.92). In all regressions, the centered stigma consciousness variable, the dummy-coded condition variable (0 = control, 1 = threat), and their interaction were entered as predictors. In the multiple regression analysis predicting OSPAN performance (Figure 1), there was a significant main effect for condition, $t(53) = -2.03$, $p = .048$ ($\beta = -.26$), such that, as expected, threat participants performed more poorly (M=16.62, SD=10.85) than did control participants (M=22.75, SD=12.46). However, this main effect was qualified by a significant interaction between condition and stigma consciousness, $t(53) = -2.11$, $p = .04$ ($\beta = -.43$). Consistent with my hypothesis, among participants in the threat condition, there was a significant negative relationship between stigma consciousness and OSPAN score ($\beta = -.57$, p=.003), whereas in the control condition, stigma consciousness was unrelated to OSPAN score ($\beta = .08$, p=.677).

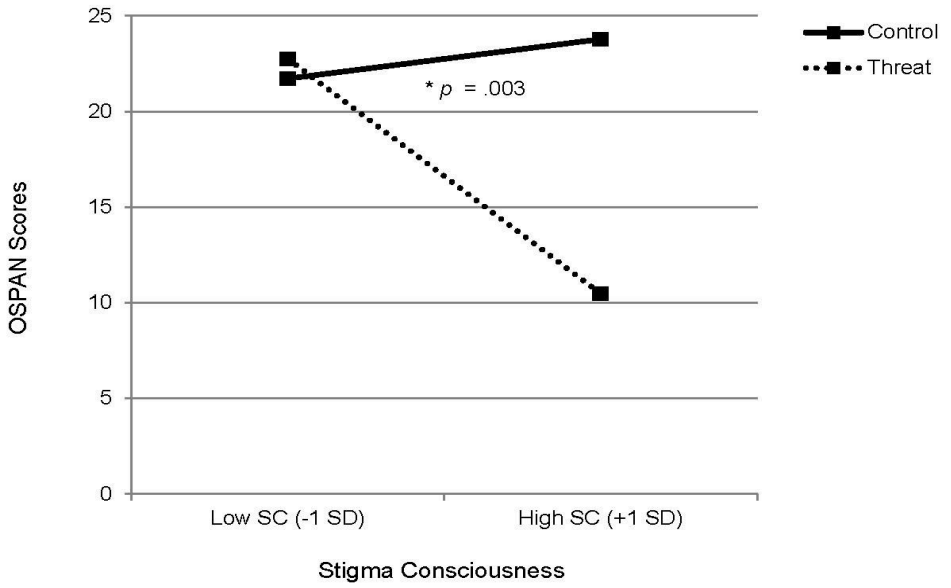

Figure 1. OSPAN performance as a function of condition and stigma consciousness.

Stereotype Suppression and Stereotype Rebound

To assess whether participants suppressed the stereotype prior to the OSPAN task, and whether these same participants then experienced post-suppressional rebound, I examined participants' reaction times to stereotypic words (relative to neutral words) for each lexical decision task (LDT). To control for individual differences in response time, participants' mean reaction time to neutral words was subtracted from their mean reaction time to stereotypic words. Thus, higher difference scores represented slower reaction times to stereotypic words relative to neutral words.

Contrary to my stereotype suppression hypothesis, neither condition, $t(53) = -.73$, $p = .472$ ($\beta = -.01$), nor stigma consciousness, $t(53) = -2.74$, $p = .52$ ($\beta = -.57$), nor the 2-way interaction, $t(53) = -2.27$, $p = .27$ ($\beta = .48$), were significant predictors of LDT1 reaction times. And, contrary to my post-suppressional rebound hypothesis, condition, $t(53) = -.10$, $p = .923$ ($\beta = -.01$), stigma consciousness, $t(53) = -.54$, $p = .59$ ($\beta = .12$), and the 2-way interaction, $t(53) = -.44$, $p = .663$ ($\beta = -.09$), were non-significant predictors of LDT2 reaction times.

To examine whether those participants with slower response times on LDT1 (and who were assumed to be suppressing) subsequently had faster responses times on LDT2 (because they were presumably experiencing rebound), I conducted a regression equation predicting LDT2 from stigma consciousness (centered), condition (dummy-coded), LDT1 (centered), and all the 2- and 3-way interactions. The only significant predictor was the stigma consciousness x LDT1 interaction, $t(46) = -2.76$, $p = .008$ ($\beta = -.52$). As illustrated in Figure 2, for participants high in stigma consciousness, there was no significant relationship between LDT1 and LDT2 reaction times ($\beta = -.13$, $p = .41$). In contrast, for

participants low in stigma consciousness, there was a strong positive relationship between LDT1 and LDT2, (β = -.69, p = .001), such that those who were quick (or slow) to react to stereotypic words before the working memory capacity task were still quick (or slow) to react to stereotypic words following the working memory capacity task.

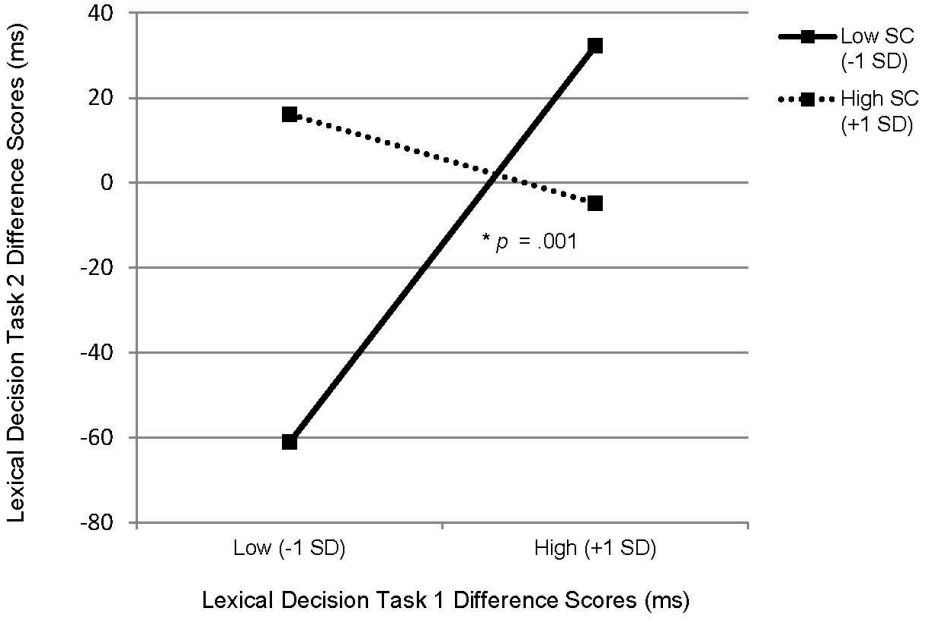

Figure 2. Lexical decision task 2 difference scores as a function of lexical decision task 1 difference scores and level of stigma consciousness. Higher difference scores indicate faster responses to stereotypic words, and lower difference scores indicate slower responses to stereotypic words.

Discussion

Stereotype Threat

Replicating past findings (Brown & Pinel, 2003), only women high in stigma consciousness in the threat condition experienced a decrement in OSPAN performance. The fact that the OSPAN performance of LSC participants did not suffer, even after the

math-gender stereotype was strongly implied and made relevant to the test situation, suggests that being low in stigma consciousness may protect against stereotype threat.

Given the overt nature of the stereotype threat manipulation (participants were told that men typically outperformed women on the OSPAN), it is unlikely that LSC women evaded stereotype threat effects because the female stereotype was not activated. Rather, it is more likely that the activated stereotype did not pose a sufficient threat to elicit stereotype threat. So, although the female stereotype was activated in LSC threat participants, it did not produce as much performance pressure as it would for those chronically aware of the math-gender stereotype. Regardless of why the salient stereotype impaired the performance of only HSC women, the findings suggest that there are, as Brown and Pinel (2003) first suggested, differences in how HSC and LSC women perform under stereotype threatening conditions.

Responses to Stereotype Threat

Domain disidentification. Unexpectedly, those in the threat condition rated performing well on the test as less important than did those in the control condition. Although this difference in test importance may account for the difference in test performance across conditions, the significant condition x stigma consciousness interaction discounts this explanation. If test unimportance was the cause of the threat participants' underperformance, then it is unlikely that there would be a distinction in the performance of HSC and LSC threat participants, let alone such a robust one. Furthermore, the test was described as a predictor of math ability and future success in only the threat condition. As such, the threat participants' lower ratings of test importance are more likely due to self-protective rationalization rather than actual indifference.

Providing further support for this rationalization account, participants responded to this question immediately following the test, while they were likely still focusing on their performance. For threat participants, who demonstrated significantly lower scores than control participants, their dissatisfactory performance may have led them to rationalize that the test was of little importance or, in other words, to salvage their self-concept by disengaging from the test. Although this instance of disengagement alone is unlikely to impact participants' future encounters with math tests, many more experiences of stereotype threat likely would. As previously discussed, repeated exposure to stereotype threat and subsequent disengagement can lead stereotype-threatened individuals to disengage from a domain entirely, as reflected by HSC students' greater academic disengagement (Brown & Lee, 2005). Therefore, these findings reinforce concerns about stereotype threat's long-term effects on academic performance.

Stereotype suppression and post-suppressional rebound. Despite indicating that they attempted to suppress stereotypic thoughts, threat participants showed no evidence of stereotype suppression, as condition was unrelated to reaction times for the first LDT. Given that threat participants did not demonstrate stereotype suppression, it follows that there was no evidence for stereotype rebound, as condition also failed to predict the reaction times for the second LDT. The most direct evidence against the stereotype suppression and post-suppressional rebound hypotheses was that the stereotype-threatened participants' reaction times to stereotypic words before and after the OSPAN were unrelated. However, I did find that, irrespective of condition, LSC participants reacted to stereotypic words consistently across the two LDTs, whereas the LDT reaction times of HSC participants were unrelated.

Although the LDT analyses did not support my hypotheses, the fact that the HSC participants responded differently to the first and second LDT suggests there may be an unidentified moderator of their reaction times. One such moderator may be individual differences in tendency to suppress the stereotype. Although there are documented instances of self-initiated stereotype suppression (Wyer et al., 1998) and subsequent post-suppressional rebound (Logel et al., 2009), research on spontaneous thought suppression, much less spontaneous thought suppression within the context of stereotype threat, is limited. It is possible that stereotype suppression is not as readily regarded as a coping strategy as previously believed. So, whereas some HSC stereotype-threatened women suppressed the stereotype, others did not, thus washing out any significant findings for the stereotype suppression and rebound hypotheses. Future researchers should assess individual differences in stereotype suppression use.

There are other possible explanations for the lexical decision task findings. For one, it is possible that the stereotypic words were not relevant to the Hamilton College community. Although the stereotypic words were taken from previous studies (Carr & Steele, 2009; Logel et al., 2009) in which the words were pilot-tested to ensure that they were evocative of the female stereotype for the researchers' student population, I did not pilot test them at Hamilton. It is possible that the math-gender stereotype manifests differently at a small liberal arts college than it does at a large, research-oriented university. If the stereotypic words represented a facet of the female stereotype different from what was activated in the stereotype-threatened participants, then the LDTs would have been incapable of detecting stereotype suppression and rebound.

Similarly, it is also possible that the math-gender stereotype is manifested differently across the different math-oriented majors. This difference is reflected by the differing gender gaps within the STEM fields. For bachelor and doctoral degrees, the gender gap in the biosciences is rapidly closing; however, female participation in mathematics, computer science, and the physical sciences remains low (National Science Foundation, 2011). Therefore, it is likely that the math-gender stereotype is more "visible" within these relatively quantitative fields than it is within the life sciences.

That is not to say that the math-gender stereotype was not activated or threatening for all threat participants during test taking. The OSPAN data strongly suggest that stereotype threat was operating on HSC threat participants regardless of major. Instead, the difference in math-gender stereotype visibility suggests that there may be more readily available cues of prejudice within mathematics, computer science, and the physical sciences, and thus women in these fields may experience stereotype threat more regularly than those in the biosciences. So, although the stereotype was still relevant and activated for HSC threat participants, stereotype-threatened STEM-students may suppress the stereotype to varying degrees given differences in their particular field.

Compare the mathematics major and the biology major, for example. Although both are well aware of the math-gender stereotype (as indicated by manipulation checks) and experience stereotype threat (as indicated by OSPAN underperformance), they may still react differently in the face of stereotype threat. Perhaps the mathematics major, now accustomed to stereotype threat after countless experiences, uses her coping strategy of choice, stereotype suppression, as she remains unaware of its negative effects. In contrast, the biology major may not have a prepared coping strategy, as she has not yet adapted to

taking tests while under stereotype threat. Yet the opposite pattern is also possible. Perhaps those unfamiliar with stereotype threat may be more likely to use stereotype suppression than more experienced stereotype-threatened students, as seasoned test takers may have developed a more sophisticated and effective means of managing stereotype threat.

However, it is also possible that experienced stereotype-threatened students still rely on stereotype suppression, but do so without experiencing post-suppressional rebound. In a study by Kelly and Kahn (1994), participants did not experience post-suppressional rebound when suppressing their own self-relevant thoughts, but did when suppressing the novel and irrelevant thought of a white bear, suggesting that people learn to suppress more successfully with experience.

Future Research

The uncertainty of whether previous experience with stereotype threat influences one's decision to suppress the stereotype (and one's success at suppressing) reflects the greater uncertainty about individual differences in stereotype suppression in general. Therefore, future research should assess whether such individual differences do exist and can start by asking participants about past self-threatening encounters. If self-reported frequency of stereotype suppression proves to be a predictor of later suppression, then stereotype threat research would greatly benefit from identifying what factors lead students to adopt this maladaptive coping strategy. Doing so will not only further our understanding of stereotype threat in general, but help to identify those especially vulnerable to the performance detriments of stereotype threat.

One possible factor may be external encouragement to use stereotype suppression. Given that stereotype suppression is a seemingly viable coping strategy, it is not difficult to imagine a well-intentioned teacher, parent, or even peer helping a distressed female math student by recommending she disregard the stereotype during test taking. Of course, without realizing, these helpers have only furthered the damaging effects of stereotype threat by compounding an already burdened working memory capacity.

If stereotype suppression is an impractical coping strategy, then what can stereotype-threatened individuals do to manage intrusive stereotypic thoughts? Logel, Iserman, Davies, Quinn, and Spencer (2009) have found that stereotype suppression may still be a feasible strategy when executed correctly. After providing participants with a neutral suppression strategy modeled after that of Wegner, Carter, Schneider, and White (1987), in which participants are instructed to substitute concerns about the test with thoughts unrelated to the test situation (e.g., a red Volkswagen), stereotype-threatened participants neither underperformed on a math test nor exhibited post-suppressional rebound. Logel and fellow researchers (2009) have also replicated these findings using more self-relevant thoughts, specifically, by asking participants to use thoughts about an important personal identity. Thus, it appears that thought-substitution, which is a more active strategy than simply avoiding distressing thoughts, improves the effectiveness of thought suppression. However, the efficacy of the thought-substitution strategy was not assessed between HSC and LSC people. It is possible, then, that thoughts about a specific personal identity, even an important one, may not be powerful enough to override the strong and widespread activation of the female stereotype in HSC participants. Future

research should address whether the thought-substitution technique is a viable option for both HSC and LSC women.

Overall Implications

Overall, the findings strongly suggest that HSC women are, in fact, more vulnerable to experiences of stereotype threat than LSC women. Given their heightened sensitivity, educators and administrators should make special effort to identify HSC women and monitor their interest in and progress through high-level math and math-related disciplines. However, the best measure against stereotype threat is to prevent it altogether. This primary prevention can be enacted through education, both at the student and administrative level. Simply informing students of stereotype threat inoculates them against its negative effects (John, Schmader, & Martens, 2005), as it allows them to attribute any test anxiety to the stereotype threat phenomenon and not their own inability. Education about stereotype threat is already on the rise; an increasing number of introductory psychology courses and textbooks now focus on the stereotype threat phenomenon (John et al., 2005). However, administrators should create programs to educate all students, regardless of academic major, of stereotype threat, as it is a widespread problem applicable to any stereotyped individual in academia.

In addition to censoring learning and testing environments of cues of prejudice, administrators should also adjust the school's perception of math ability from permanent and inheritable to malleable and incremental, as such attitudinal changes will foster a greater sense of belonging among female students within the math community, which in turn encourages these women to continue with math and math-related disciplines (Good, Rattan, & Dweck, 2012).

Conclusion

Together, the results of the present study demonstrate that stereotype threat negatively affects women's math performance, and that those high in stigma consciousness are susceptible to these effects. Although the present study did not find evidence of spontaneous stereotype suppression and post-suppressional rebound, future studies should continue with stereotype suppression research by addressing whether there are individual differences in stereotype suppression use. Furthering stereotype suppression research will not only deepen researchers' understanding of how stereotype threat operates, but will allow those beyond the laboratory to implement more meaningful and effective strategies to mitigate stereotype threat.

References

Aronson, J., Lustina, M. J., Good, C., Keough, K., Steele, C. M., & Brown, J. (1999). When white men can't do math: Necessary and sufficient factors in stereotype threat. *Journal of Experimental Social Psychology, 35*(1), 29-46. doi: 10.1006/jesp.1998.1371

Bargh, J. A., Chen, M., & Burrows, L. (1996). Automaticity of social behavior: Direct effects of trait construct and stereotype activation on action. *Journal of Personality and Social Psychology, 71*(2), 230-244. doi: 10.1037/0022-3514.71.2.230

Beede, D., Julian, T., Langdon, D., McKittrick, G., Khan, B., & Doms, M. (2011). Women in STEM: a gender gap to innovation. *U.S. Department of Commerce: Economics and Statistics Administration*. Retrieved from http://www.esa.doc.gov/Reports/women-stem-gender-gap-innovation

Beilock, S. L., & R. (2007). Stereotype threat and working memory: Mechanisms, alleviation, and spillover. *Journal of Experimental Psychology: General, 136*, 256-276. doi: 10.1037/0096-3445.136.2.256

Beilock, S. L., Jellison, W. A., Rydell, R. J., McConnell, A. R., & Carr, T. H. (2006). On the causal mechanisms of stereotype threat: Can skills that don't rely heavily on working memory still be threatened? *Personality and Social Psychology Bulletin, 32*(8), 1059-1071. doi: 10.1177/0146167206288489

Ben-Zeev, T., Fein, S., & Inzlicht, M. (2005). Arousal and stereotype threat. *Journal of Experimental Social Psychology, 41*(2), 174-181. doi: 10.1016/j.jesp.2003.11.007

Borton, J. L. S., Reiner, D. R., Vazquez, E. V., Ruddiman, J. F., & Anglin, S. M. (2011). Does suppressing the thought of a self-relevant stigma affect interpersonal interaction? *The Journal of Social Psychology, 151*(3), 240-256. doi: 10.1080/00224541003599050

Brown, R. P., & Lee, M. N. (2005). Stigma consciousness and the race gap in college academic achievement. *Self and Identity, 4*(2), 149-157. doi: 10.1080/13576500444000227

Brown, R. P., & Pinel, E. C. (2003). Stigma on my mind: Individual differences in the experience of stereotype threat. *Journal of Experimental Social Psychology, 39*(6), 626-633. doi: 10.1016/S0022-1031(03)00039-8

Cadinu, M., & Maass, A., Rosabianca, A., Kiesner, J. (2005). Why do women underperform under stereotype threat? evidence for the role of negative thinking. *Psychological Science, 16*, 572-578. doi: 10.1111/j.0956-7976.2005.01577.x

Ceci, S. J., Williams, W. M., & Barnett, S. M. (2009). Women's underrepresentation in science: Sociocultural and biological considerations. *Psychological Bulletin, 135*(2), 218-261. doi: 10.1037/a0014412; 10.1037/a0014412.supp (Supplemental)

Carr, P. B., & Steele, C. M. (2009). Stereotype threat and inflexible perseverance in problem solving. *Journal of Experimental Social Psychology, 45*(4), 853-859. doi: 10.1016/j.jesp.2009.03.003

Cohen, J., MacWhinney, B., Flatt, M., & Provost, J. (1993). Psyscope: An interactive graphic design system for designing and controlling experiments in the psychology laboratory using Macintosh computers. *Behavior Research Methods, Instruments, and Computers, 25*, 257-271.

Croizet, J., & Claire, T. (1998). Extending the concept of stereotype and threat to social class: The intellectual underperformance of students from low socioeconomic backgrounds. *Personality and Social Psychology Bulletin, 24*(6), 588-594. doi: 10.1177/0146167298246003

Croizet, J. C., Després, G., Gauzins, M. E., Huguet, P., Leyens, J. P., & Méot, A. (2004). Stereotype threat undermines intellectual performance by triggering a disruptive mental load. *Personality and Social Psychology Bulletin*, *30*(6), 721-731. doi: 10.1177/0146167204263961

Danso, H. A., & Esses, V. M. (2001). Black experimenters and the intellectual test performance of white participants: The tables are turned. *Journal of Experimental Social Psychology, 37*(2), 158-165. doi: 10.1006/jesp.2000.1444

Davies, P. G., Spencer, S. J., Quinn, D. M., & Gerhardstein, R. (2002). Consuming images: How television commercials that elicit stereotype threat can restrain women academically and professionally. *Personality and Social Psychology Bulletin, 28*(12), 1615-1628. doi: 10.1177/014616702237644

Dickler, J., Lee, C. A., Swiatek, G. (2011). *Best jobs for fast growth.* Retrieved November 4, 2012, http://money.cnn.com/magazines/moneymag/best-jobs/2011/fast-growing-jobs/1.html.

Dijksterhuis , A. Bargh , J. A. (2001). The perception-behavior expressway. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 33, pp. 1-40). San Diego, CA: Academic Press.

Dijksterhuis, A., & van Knippenberg, A. (1998). The relation between perception and behavior, or how to win a game of trivial pursuit. *Journal of Personality and Social Psychology, 74*(4), 865-877. doi: 10.1037/0022-3514.74.4.865

Gonzales, P. M., Blanton, H., & Williams, K. J. (2002). The effects of stereotype threat and double-minority status on the test performance of Latino women. *Personality and Social Psychology Bulletin, 28*(5), 659-670. doi: 10.1177/0146167202288010

Good, C., Aronson, J., & Harder, J. A. (2008). Problems in the pipeline: Stereotype threat and women's achievement in high-level math courses. *Journal of Applied Developmental Psychology, 29*(1), 17-28. doi: 10.1016/j.appdev.2007.10.004

Good, C., Rattan, A., & Dweck, C. S. (2012). Why do women opt out? Sense of belonging and women's representation in mathematics. *Journal of Personality and Social Psychology*, *102*(4), 700.

Inzlicht, M., & Ben-Zeev, T. (2000). A threatening intellectual environment: Why females are susceptible to experiencing problem-solving deficits in the presence of males. *Psychological Science, 11*(5), 365-371. doi: 10.1111/1467-9280.00272

Johns, M., Schmader, T., & Martens, A. (2005). Knowing Is Half the Battle Teaching Stereotype Threat as a Means of Improving Women's Math Performance. *Psychological Science*, *16*(3), 175-179.

Kaiser, C. R., Vick, S. B., & Major, B. (2006). Prejudice expectations moderate preconscious attention to cues that are threatening to social identity. *Psychological Science, 17*(4), 332-338. doi: 10.1111/j.1467-9280.2006.01707.x

Kelly, A. E., & Kahn, J. H. (1994). Effects of suppression of personal intrusive thoughts. *Journal of Personality and Social Psychology*, *66*(6), 998.

Kiefer, A. K., & Sekaquaptewa, D. (2007). Implicit stereotypes, gender identification, and math-related outcomes: A prospective study of female college students. *Psychological Science, 18*(1), 13-18. doi: 10.1111/j.1467-9280.2007.01841.x

Langdon, D., McKittrick, G., Beede, D., Khan, B., & Doms, M. (2011). STEM: Good jobs now and for the future. *U.S. Department of Commerce: Economics and Statistics Administration*. Retrieved from http://www.esa.doc.gov/sites/default/files/reports/documents/stemfinaljuly14.pdf

Lindberg, S. M., Hyde, J. S., Petersen, J. L., & Linn, M. C. (2010). New trends in gender and mathematics performance: A meta-analysis. *Psychological Bulletin, 136*(6), 1123-1135. doi: 10.1037/a0021276; 10.1037/a0021276.supp (Supplemental)

Logel, C., Iserman, E. C., Davies, P. G., Quinn, D. M., & Spencer, S. J. (2009). The perils of double consciousness: The role of thought suppression in stereotype threat. *Journal of Experimental Social Psychology, 45*(2), 299-312. doi: 10.1016/j.jesp.2008.07.016

Logel, C., Walton, G. M., Spencer, S. J., Iserman, E. C., von Hippel, W., & Bell, A. E. (2009). Interacting with sexist men triggers social identity threat among female engineers. *Journal of Personality and Social Psychology, 96*(6), 1089-1103. doi: 10.1037/a0015703

Macrae, C. N., Bodenhausen, G. V., Milne, A. B., & Jetten, J. (1994). Out of mind but back in sight: Stereotypes on the rebound. *Journal of Personality and Social Psychology*, *67*(5), 808. doi: 10.1037/0022-3514.67.5.808

Major, B., Spencer, S., Schmader, T., Wolfe, C., & Crocker, J. (1998). Coping with negative stereotypes about intellectual performance: The role of psychological

disengagement. *Personality and Social Psychology Bulletin*, *24*(1), 34-50. doi: 10.1177/0146167298241003

National Science Foundation. (2011). *Women, minorities, and persons with disabilities in science and engineering: 2011.* Retrieved November 4, 2012, http://www.nsf.gov/statistics/wmpd/pdf/nsf11309.pdf

O'Brien, L. T., & Crandall, C. S. (2003). Stereotype threat and arousal: Effects on women's math performance. *Personality and Social Psychology Bulletin, 29*(6), 782-789. doi: 10.1177/0146167203029006010

Osborne, J. W. (1997). Race and academic disidentification. *Journal of Educational Psychology*, *89*(4), 728. doi: 10.1037/0022-0663.89.4.728

Pinel, E. C. (1999). Stigma consciousness: The psychological legacy of social stereotypes. *Journal of Personality and Social Psychology, 76*(1), 114-128. doi: 10.1037/0022-3514.76.1.114

Pinel, E. C. (2002). Stigma consciousness in intergroup contexts: The power of conviction. *Journal of Experimental Social Psychology, 38*(2), 178-185. doi: 10.1006/jesp.2001.1498

Pinel, E. C. (2004). You're just saying that because I'm a woman: Stigma consciousness and attributions to discrimination. *Self and Identity, 3*(1), 39-51. doi: 10.1080/13576500342000031

Pinel, E. C., & Paulin, N. (2005). Stigma consciousness at work. *Basic and Applied Social Psychology, 27*(4), 345-352. doi: 10.1207/s15324834basp2704_7

Pinel, E. C., Warner, L. R., & Chua, P. (2005). Getting there is only half the battle: Stigma consciousness and maintaining diversity in higher education. *Journal of Social Issues, 61*(3), 481-506. doi: 10.1111/j.1540-4560.2005.00417.x

Schmader, T. (2002). Gender identification moderates stereotype threat effects on Women's math performance. *Journal of Experimental Social Psychology, 38*, 194-201. doi: 10.1006/jesp.2001.1500

Schmader, T., & Johns, M. (2003). Converging evidence that stereotype threat reduces working memory capacity. *Journal of Personality and Social Psychology, 85*(3), 440-452. doi: 10.1037/0022-3514.85.3.440

Schmader, T., Johns, M., & Forbes, C. (2008). An integrated process model of stereotype threat effects on performance. *Psychological Review, 115*(2), 336-356. doi: 10.1037/0033-295X.115.2.336

Sekaquaptewa, D., & Thompson, M. (2003). Solo status, stereotype threat, and performance expectancies: Their effects on women's performance. *Journal of Experimental Social Psychology, 39*(1), 68-74. doi: 10.1016/S0022-1031(02)00508-5

Spencer, S. J., Steele, C. M., & Quinn, D. M. (1999). Stereotype threat and women's math performance. *Journal of Experimental Social Psychology, 35*(1), 4-28. doi: 10.1006/jesp.1998.1373

Steele, C. M. (1997). A threat in the air: How stereotypes shape intellectual identity and performance. *American Psychologist, 52*(6), 613-629. doi: 10.1037/0003-066X.52.6.613

Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology, 69*(5), 797-811. doi: 10.1037/0022-3514.69.5.797

Stone, J. (2002). Battling doubt by avoiding practice: The effects of stereotype threat on self-handicapping in white athletes. *Personality and Social Psychology Bulletin, 28*(12), 1667-1678. doi: 10.1177/014616702237648

Stone, J., Lynch, C. I., Sjomeling, M., & Darley, J. M. (1999). Stereotype threat effects on Black and White athletic performance. *Journal of Personality and Social Psychology; Journal of Personality and Social Psychology*, *77*(6), 1213. doi: 10.1037/0022-3514.77.6.1213

Stone, J., & McWhinnie, C. (2008). Evidence that blatant versus subtle stereotype threat cues impact performance through dual processes. *Journal of Experimental Social Psychology, 44*(2), 445-452. doi: 10.1016/j.jesp.2007.02.006

Turner, M. L., & Engle, R. W. (1989). Is working memory capacity task dependent? *Journal of Memory and Language, 28*(2), 127-154. doi: 10.1016/0749-596X(89)90040-5

Wegner, D. M. (1994). Ironic processes of mental control. *Psychological Review, 101*(1), 34-52. doi: 10.1037/0033-295X.101.1.34

Wegner, D. M., & Erber, R. (1992). The hyperaccessibility of suppressed thoughts. *Journal of Personality and Social Psychology, 63*(6), 903-912. doi: 10.1037/0022-3514.63.6.903

Wegner, D. M., & S. (1987). Paradoxical effects of thought suppression. *Journal of Personality and Social Psychology, 53*, 5-13. Retrieved from

Wheeler, S. C., & Petty, R. E. (2001). The effects of stereotype activation on behavior: A review of possible mechanisms. *Psychological Bulletin, 127*(6), 797-826. doi: 10.1037/0033-2909.127.6.797

Wheeler, S. C., DeMarree, K. G., & Petty, R. E. (2007). Understanding the role of the self in prime-to-behavior effects: The active-self account. *Personality and Social Psychology Review*, *11*(3), 234-261. doi: 10.1177/1088868307302223

Wout, D., Danso, H., Jackson, J., & Spencer, S. (2008). The many faces of stereotype threat: Group- and self-threat. *Journal of Experimental Social Psychology, 44*(3), 792-799. doi: 10.1016/j.jesp.2007.07.005

Wyer, N. A., Sherman, J. W., & Stroessner, S. J. (1998). The spontaneous suppression of racial stereotypes. *Social Cognition, 16*(3), 340-352. doi: 10.1521/soco.1998.16.3.340