12-2018

The Impact of Stereotype Threat on Object Location Memory

Samantha A. Boomgarden

THE IMPACT OF STEREOTYPE THREAT ON OBJECT LOCATION MEMORY

An Honors Thesis submitted to the

Department of Psychology
College of Sciences
And
Honors College
Eastern Illinois University

In partial fulfillment of the requirements
Of the Departmental Honors Program
For the degree of

BACHELOR OF ARTS

In the Department of Psychology

December, 2018

By

Samantha A Boomgarden

Successfully defended this 6th day of December, 2018

Thesis Advisor: _____

Committee Member: _____

Honors Coordinator: _____

Department Chair: _____

The Impact of Stereotype Threat on Object Location Memory

Samantha Boomgarden

Eastern Illinois University

Table of Contents

Abstract

Gender differences have been identified in many tasks, and the male advantage in spatial skills has been well studied and is thought to be robust, especially on mental rotation and spatial perception tasks (e.g., Doyle & Voyer, 2016; Linn & Petersen, 2016; Pansu et al., 2016; Thompson & Voyer, 2014). However, women have been found to do better on tasks that require memorization of where objects are located in the environment (i.e., object location memory tasks; Voyer, Postma, Brake, & Imperato-McGinley, 2007). The purpose of this study was to examine how stereotype threat, elicited in women, would affect their performance in an object location memory task. Mediating factors such as gender identification and state anxiety were also analyzed. Participants were randomly assigned either to a group that would be presented with the stereotype threat or one that did not receive the threat. It was hypothesized that those in the threat group would have poorer performance compared to those not threatened. An independent samples t-test was used to analyze the performance on an object location memory task. No significant difference on object location memory scores was found between the participants who were presented with the stereotype threat and those who were not. However, a significant negative correlation was found between participant's state anxiety and their gender identification in the threat group.

Keywords: stereotype threat, gender identification, state anxiety

Introduction

Gender differences have been identified in many tasks, some with a female advantage, such as reading performance and emotion recognition, and some with a male advantage, such as mathematical assessments and spatial skills (e.g., Doyle & Voyer, 2016; Linn & Petersen, 2016; Pansu et al., 2016; Thompson & Voyer, 2014). The male superiority in spatial skills, in particular, has been well-studied and thought to be robust, especially on mental rotation and spatial perception tasks, (Linn & Petersen, 2016). Spatial skills have been tested in various ways, such as with spatial orientation tasks, spatial visualization tasks, visuo-motor integration tasks, and even jig-saw puzzles. There is at least one spatial task where women have been found to outperform men, and that is on tasks that require memorization of where objects are located in the environment (i.e., object location memory tasks; Voyer et al., 2007).

Object location memory is a skill that we seem to begin to use at a young age; for example, the simple child's game of matching cards after they have been flipped (e.g., as in the game "Concentration") can be enjoyable for children. Among adults, when it comes to object location memory tasks, women actually tend to perform better than men (Voyer et al., 2007). One explanation for this gender difference has been presented in the context of the foraging hypothesis, an evolutionary perspective, proposed by Silverman and Eals (1992), which theorizes that women's spatial processing developed to allow them to more efficiently remember the location of plants from season to season so they could fulfill their gathering role. A more recent, broader study supported this female advantage on location memory based on data gathered from 40 countries (Silverman, Choi, & Peters, 2007). This skill may persist among women, perhaps because it is may be used every day, for example, in remembering where an item is located in a store, where a child is playing in a playground, or even in the broader environment around us. Object location memory may well be an important and useful skill. But, the debate about

whether it is a culturally-specific developed skill remains (Hoffman, Gneezy, & List, 2011); and that debate may persist because the parameters that can impact this skill are not completely well-understood and research questions remain.

Individuals' expectations about how they are likely to perform on tasks has been found to affect how they actually perform on those tasks, including memory tasks (Mazerolle, Régner, Rigalleau, & Huguet, 2015). So, it is likely that performance on object location memory tasks could be impacted by the personal expectations that individuals have developed about such tasks. One example of how an individual's expectation about task performance might affect subsequent performance on the task by the individual is stereotype threat.

Stereotype threat has been identified as a way to impact performance on various memory and skills tasks (e.g., Huber, Seitchik, Brown, Sternad, & Harkins, 2015; Mazerolle et al., 2015), although its impact on object location memory is unclear. Stereotype threat is the phenomena in which an individual's performance is negatively impacted by the presence of a negative stereotype about one of their social identities (e.g., race, gender) (e.g., Spencer, Steele, & Quinn, 1999; Steele & Aronson, 1995); for example, a Black student's performance on an assessment is likely to diminish if the student is told that the assessment is indicative of intelligence, because of the societally instilled stereotype that Black students are not as intelligent as White students. However, this diminished performance may have little to do with the Black student's actual ability, because, if given that test in an environment without the threat, the student may perform just as well, or better, than their White counterparts (Steele & Aronson, 1995).

The impact of stereotype threat has also been observed between genders. For instance, Spencer, Steele, and Quinn (1999) found that presenting a task as one that required quantitative skills resulted in the lower performance of female students—presumably, because women may

tend to more readily accept the societal view that they are less quantitatively skilled. Similar results have been found in the performance of female students on GRE quantitative tests in the presence of a stereotype threat (Jamieson & Harkins, 2007). In fact, a meta-analysis of stereotype manipulation on math performance found that the performance of women was negatively impacted by the presence of a stereotype threat in math tasks (Doyle & Voyer, 2016). Thus, research on the impact of stereotype threat in quantitative domains has been most documented. But, it also seems to exist in other domains, such as spatial skills. For example, Tarampi, Heydari, and Hegarty (2016) framed a task as spatial and found that female performance decreased with the presence of the societal stereotype about women's spatial abilities.

Still, it remains unclear what the impact of a stereotype threat on women's object location memory performance would be. The purpose of the current study was to examine how the presence of a stereotype threat, introduced by informing women that they would be performing a spatial task and that men do tend to perform better on spatial tasks, would affect performance on the task. Because activation of a stereotype threat has been shown to increase anxiety or awareness (Chung, Ehrhart, Ehrhart, Hattrup, & Solamon, 2010), an adapted version of the Spielberger State-Trait Anxiety Inventory (STAI) was also administered; and because how much a women identifies with her gender might also affect the impact of the stereotype threat, a gender identification questionnaire (GIQ), adapted from Multigroup Ethnic Identity Scale (MEIM), was also administered to quantify the strength of each woman's gender identification (Phinney, 1992).

Literature Review

Stereotype Threat

Stereotype threat was first coined by Steele and Aronson (1995), who found that Black students' performance diminished when a test they were administered was framed as diagnostic

of intelligence; Steele and Aronson hypothesized that framing the test as an intelligence test activated a stereotype regarding the intelligence of Black Americans. This was not the first time the effects of stereotype threat had been investigated, however. In 1964, Katz, Epps, and Axelson examined the performance of Black participants in all-Black or all-White environments. When the participants were under a low threat environment—meaning the task was easy with little stress—the Black students performed better when they were placed in the all-White environment. Conversely, when the Black students were in a high threat environment, those in the all-White environment performed more poorly than those in the all-Black environment. It was hypothesized that this interaction was caused by the Black participant being motivated by societal stressors of being in competition with the other white participants. This may not be precisely the stereotype threat that we discuss now; however, this impact on performance is similar. When the task was low threat, the stress was just enough to improve performance, but when the threat was higher, the stress increased enough to impair performance. These results could be due to one of the potential mechanisms for the impact of stereotype threat: the mere effort account.

The mere effort account hypothesizes that if an individual has strategies that would lead to them performing better on a task prior to the presentation of the threat, they will be less impacted by the threat—in fact, their performance may even improve (Jamieson & Harkins, 2007). The mere effort account claims that the presence of a stereotype threat will activate the prepotent, or most likely, response. The added pressure of being evaluated in the context of a negative stereotype activates this response. When a stereotype threat is presented, it creates a situation where the individual is concerned with confirming the negative stereotype about their social identity. If the individual has strategies that would lead to them performing worse, the

stereotype threat will negatively impact their performance. However, if the individual has strategies that would make the task easier, the threat may improve their performance.

Spencer, Steele, and Quinn (1999) applied previously race-based research to investigate gender differences, most specifically in quantitative assessments. When a quantitative task based on the advanced GRE quantitative section was framed as producing gender differences, women significantly underperformed compared to the men in the study. However, when the same task was described as not producing a gender difference in results, women performed equally as well as the men. They hypothesized that this decreased performance is because women in the former group experienced a stereotype threat in regards to their math ability.

Investigators have continued exploring how stereotype has impacted gender differences on many different types of tasks. The effects of different types of anxiety on female students' performance in a statistics class was examined (Kapitanoff & Pandey, 2017). When women endorsed the stereotype regarding their own gender's performance on mathematical assessments, they reported higher anxiety, both about the specific class they were in and about math in general. Other studies have explored the impact of stereotype threat on athletic ability, interest in science, technology, engineering, and mathematics (STEM) fields, as well as continued research about performance in quantitative domains (Casad, Hale, & Wachs, 2017; Hively & El-Alayli, 2014; Shapiro & Williams, 2012). Hively and El-Alayli found that stereotype threat can negatively impact women's athletic performance when a task was more difficult; however, when the task was easy, the participants were not impacted by the stereotype threat. Shapiro and Williams demonstrated that there are multiple sources that can impact girls' interest in mathematical fields: parents' and teachers' gender-related math attitudes and knowledge of

stereotypes regarding gender differences in math performance. In sum, stereotype threat effects regarding female performance can be found in many different contexts

Spatial Skills and Stereotype Threat

The distinction among the various types of spatial skills task is seldom presented; indeed an online browser search for "are men better at spatial tasks?" brings up an array of references affirming that men have better spatial skills than women; it is only upon further detailed reading that authors might distinguish among specific spatial skills. Thus, the use of an object location memory task presents a unique opportunity to further investigate how stereotype threats may impact women's task performance because it is likely that non-researchers may already have a stereotype about men's superiority in spatial skills without any notion of outcomes for specific tasks. Although women tend to perform better on object location memory, the general belief that men are better at spatial tasks allows for the manipulation of a stereotype threat in an experimental setting.

Mental Rotation Tasks. Mental rotation, has an advantage in the opposite direction—men significantly outperform women (Boone & Hegarty, 2017). It has been hypothesized that this advantage comes from the application of solution strategies to solve various mental rotation tasks. When a mental rotation task is easier (i.e. the angular disparities are below 90^{o}), individuals will rotate the objects to see if they are the same; however if the angular disparity is above 90^{o}, then individuals typically switch to another strategy to solve the problem. On these more difficult problems, Boone and Hegarty hypothesized that women perform more poorly because they focus on the details of the shape (i.e. how many blocks are present), rather than on the global shape like men do. However, when women were made aware of a specific strategy that was beneficial to solving the mental rotation task, their performance improved.

McGlone and Aronson (2006) examined the impact of stereotype threat on a mental rotation task. Female participants who were reminded of their identity as high-achieving college students performed better than female participants primed to consider their gender. The impact of the stereotype threat indicates that stereotype threat can indeed impact performance on at least one spatial task.

Other Spatial Skills Tasks. Framing a spatial task as a test of spatial abilities has been shown to lead to similar better male performance outcomes. Tarampi, Heydari, and Hegarty (2016) framed the assessment as a spatial task and tested both male and female participants. In the spatial condition, the researchers found that men had an advantage; however, when they included human figures and framed the same task as social, women actually had the advantage and performed better than men. Tarampi et al. suggested that the improvement in scores from women in the social condition resulted from both the addition of the human figures and the new perspective on the task. They also suggested that stereotype threat was responsible, in part, for the diminished performance of women in the spatial condition.

Doyle and Voyer (2016) conducted a meta-analysis on stereotype threat on both spatial and math tasks. The researchers found significant results of a stereotype threat's impact on female performance when the task was math-related, however, they did not find the same significance with spatial tasks. This was hypothesized to be due to a lower societal awareness of the stereotype about male performance on spatial tasks. However, the tasks identified in the meta-analysis were all mental rotation spatial tasks, with the male advantage, rather than object location memory tasks.

Object Location Memory

One skill where little stereotype threat specific research has been focused is object location memory (OLM), a task coined by Silverman and Eals in 1992. Object location memory is a spatial skill that some research indicates brings about a gender difference where women perform better (Voyer 2007; Silverman and Eals, 1992; Eals & Silverman, 1994; Tottenham, Saucier, Elias, & Gutwin, 2003; Spiers, Sakamoto, Elliott, & Baumann, 2008). Evolutionary psychologists say that this skill harkens back to our hunter-gatherer days. It was more beneficial for women to remember where different objects like fruits and berries were located in the woods, therefore, women perform better now on object location tasks. This theory is supported by other evidence of evolutionarily selected traits beneficial for the "gathering" role found within the female population such as fewer instances of color-blindness, higher color discrimination, and higher accuracy in smell recognition (Ecuyer-Dab & Robert, 2007).

The hunter-gatherer hypothesis was further examined by having male and female participants perform an object location memory assessment with plants (Neave, Hamilton, Hutton, Tildesley, & Pickering, 2005). Previous studies, such as those performed by Silverman and Eals (1992, 1994), had only used drawings of objects such as furniture and tools. However, Neave and colleagues (2005) argued that the use of these objects did not support the division of labor hypothesis because the stimuli was not ecologically valid; instead, the stimuli should better represent the objects women would have used these spatial skills to find, for example, plants. The expected female advantage in object location memory was still found with the use of these ecologically valid stimuli, offering further support for Silverman and Eal's (1992) initial hypothesis.

This female advantage is also found in other contexts, such as in a grocery store (Spiers et al., 2008). A virtual reality spatial object location test was devised to investigate whether object location memory still had a female advantage within a functional task. Male and female participants simulated grocery shopping by using a joystick to control the program as they completed the task of picking up 12 specified items in the store. After they picked up the items, they reported where they found the 12 items on a 2D map, as well as the location of 4 incidental items. The 4 incidental items were included in the second part of the task because it has been hypothesized that the female advantage in object location memory is more robust when the learning is incidental rather than direct (McGivern et al., 1998). In the grocery shopping assessment, women outperformed men when asked to recall the locations of the 16 items. The researchers hypothesized that this advantage could be due to cultural expectations regarding gender roles—women, societally, are more expected to do the grocery shopping—so the participants also completed a 2D object location memory task. Women again outperformed men, therefore, the researchers suggested that the female advantage in the grocery shopping task was due to object location memory differences (Neave et al., 2005).

Mediating Factors: Anxiety & Gender Identification

There are factors that can mediate the impact of a stereotype threat. Anxiety and identification with the stereotype identity are two such mediators (Chung, Ehrhart, Ehrhart, et al., 2010; Deaux et al., 2007).

Chung et al. (2010) investigated whether the relationship between perception of a stereotype threat and performance on an assessment would be mediated by both state anxiety, specific self-efficacy, and ethnic identity. The researchers examined individuals' performances on a job assessment in the real world. Participants completed the assessment for their profession

and then completed various self-report measures to quantify their perception of a stereotype threat, state anxiety, and self-efficacy all while completing the assessment, as well as their ethnic identity. Results from this study indicated that higher state anxiety was correlated with lower self-efficacy, which in turn, was correlated to a lower performance on the assessment. Results regarding the impact of ethnic identity on performance were unclear, but they did find that a relationship between high ethnic identity and increased perception of a stereotype threat was present. Most noticeably, however, was that higher state anxiety was indeed found to be associated with lower scores.

One measure of anxiety is the Spielberger State-Trait Anxiety inventory (1983). This can measure both state and trait anxiety through a series of questions that ask participants to indicate how often they feel a certain way, both generally and in that moment. The questions that ask about feelings generally are designed to measure trait anxiety, or how anxious a person is in their everyday life. The questions that ask more specifically about the moment are designed to measure state anxiety, or how anxious a person reports feeling in the moment. For the purposes of this study, only state anxiety was necessary for analysis.

Another important factor in the potential impact a stereotype threat can have is the individual's identification with the threatened social identity. Deaux and colleagues (2007) presented a stereotype threat to first-generation immigrants from the West Indies and found something called *stereotype lift*—an improvement in performance when presented with a stereotype about an out-group. The participants were told that they were completing a task that Black American individuals performed worse on. Since the first-generation immigrants did not see themselves as the group the stereotype was presented about, their performance was actually improved. However, second generation immigrants were found to be negatively impacted by the

same stereotype threat because the threat impacted what they saw as their social identity (Deaux et al., 2007). The second-generation immigrants were negatively impacted by the stereotype threat because they identified with the social identity being stereotyped. Steele (1997) proposed that these "domains" were personal identities that an individual sees as evaluatively accountable, that is, they are impacted by the judgement of their performance within the domain. This is particularly evident in academic achievement: when a student identifies with this domain of academia it is something that they define themselves with. Their self-definition is more impacted by their performance within this domain. These domains could be other aspects of a social self besides academic achievement such as gender or ethnicity.

In the present study, the participants' identification with the female gender was quantified using an adapted version of the multigroup ethnic identity scale (Phinney, 1992). Instead of asking about ethnic identity, the different questions asked about gender identity and the individual's sense of belonging to their own gender group, which, for the purpose of this study, was always female.

Current Study

Hypothesis 1: The presentation of a stereotype threat concerning women's performance on a spatial task should result in diminished performance, even though there is a pre-existing female advantage on OLM tasks. Compared to a no-threat control group, women in a stereotype threat group were hypothesized to (a) perform more poorly on the OLM task, and (b) have higher state anxiety scores.

Hypothesis 2: Gender identification was also hypothesized to be negatively correlated with OLM task performance. That is, women who identify more strongly with their gender would perform worse on the OLM task. However, this correlation should be observed for the threat condition but not the no-threat condition.

Hypothesis 3: State anxiety was hypothesized to be negatively correlated with OLM task performance. Women who show higher state anxiety were hypothesized to perform worse on the OLM task.

Hypothesis 4: When under stereotype threat, high levels of gender identification were hypothesized to be positively correlated with higher state anxiety scores.

Method

Participants

An a priori test was performed to determine the needed sample size for the independent t-test with a moderate effect size (0.5) and a power of 0.8. The test indicated that 51 participants per condition would be needed.

Participants were recruited using the university's research participation program online software. Therefore, all participants were from Eastern Illinois University and a majority were in an introductory psychology class. The mean age was 19.1. A prescreen restriction was enabled so that only female students could sign up for the study. Participants received 1 credit of participation in research, which is either required or is extra credit for their psychology course.

In total, 105 participants were recruited. They were randomly assigned to one of two conditions: either they were presented with the stereotype threat (experimental) or they weren't (control). There were 49 participants in the experimental condition and 50 participants in the control condition whose responses were included in the final data analysis. Those who did not complete the whole survey (N=3) and those who did not pass the manipulation checks (N=3) were excluded from final data analysis.

Materials

Spielberger State-Trait Anxiety Inventory (STAI). An adapted version of the Spielberger State-Trait Anxiety Inventory (1983) was developed. This version was much shorter (only 10 trait questions and 10 state questions). The Trait answers ranged from "Not at all" to "All the time"; however, the trait portion was not scored or used for analysis. The State answers ranged from "Not at All" to "Very Much So" for the State portion, which specifically measured their self-reported anxiety while taking the assessment. A sliding scale was used to indicate how often they felt the way each question indicated. This was scored from 0-100, with 0 indicating "Not At All" and 100 indicating "Very Much So". Questions that asked about non-anxious feelings, i.e. "I feel calm", were reverse scored, so that high scores on this measure indicated high levels of self-reported state anxiety. The average score was calculated with all the questions for state anxiety was used for analysis. Scores could range from 0-100, and the mean score reported was 31.78, with a standard deviation of 16.61. The Cronbach's alpha for the adapted State scale was .906. (See Appendix A).

Gender Identity. A version of the multigroup ethnic identity scale was adapted to quantify the participant's identification with their own gender (GIQ) (Phinney, 1992). Instead of inquiring about an ethnic group, the adapted version asked about gender. To indicate how much they agreed with a statement, there were choices ranging from strongly agree (scored as 5) to strongly disagree (scored as 1). The average of the answers from the 10 GIQ items was recorded and used as the GIQ score in analysis. A high score on this measure indicated high levels of self-reported gender identification. Scores could range from 1-5, and the mean score was 3.99, with a standard deviation of .49. The Cronbach's alpha for this scale was .704. (See Appendix B).

OLM Task. An OLM task was created based on those developed by Silverman and Eals (1992). Three arrays of 42 images each—the encoding arrays—were developed. Each array was used in a separate test trial, for a total of 3 test trials. A practice trial was also developed with only 20 images. For each trial an encoding array was programmed to be presented for one minute. A distractor task, which involved sorting objects or concepts, i.e. sorting a list of animals into amphibian, insect, bird, or reptile groups, appeared immediately after the 42 image encoding array, for another minute. Then, a test array with same 42 images was presented, but with 21 objects (10 objects for the practice trial) moved to a different location on the array, again for one minute. The task required participants to identify which objects in the test array had moved from the first array; correct identification of objects that had moved were scored as hits; incorrect identification of objects as having moved were scored as false alarms. Images for the arrays were obtained from a Google search for images under general license and placed in an approximate grid formation—6 across and 7 down. (See Appendix C for the first trial images).

Instructional Videos. Three instructional videos were created for the participants to watch. (1) The first was a brief instructional video that explained the practice testing procedure: participants would have 30 seconds to study an array of images (this array only had 20 objects); then they would have another 30 seconds to complete a sorting task; and then they would be shown another array with the same objects, however some of the objects would have moved. They would, again, have 30 seconds to click on the objects that had moved in the second array. (2) The second video explained that instead of 30 seconds for each portion of the trial, participants would now have a minute to complete each task and they would have 42 objects instead of 20. (3) The third video was different for the threat and no-threat groups. In this video, those in the threat group were told that men tended to perform better on these spatial

assessments, and the current study was examining gender differences. The no-threat group received a re-iteration of the initial instructions.

Additional Questions. Participants were asked a series of questions as attention checks after watching the videos. After the second video, these questions asked how many trials they had to complete, which array would have the objects moved, and how they should indicate which objects moved. After the third video, these questions asked what kind of assessment they would be completing, how many trials were left, and on which array they needed to click on the objects. Participants in the threat group had an additional question that asked them to indicate which gender performed better on this type of assessment as a manipulation check. After completing the 3 trials of the OLM task, participants in the threat group had one additional question: "What did you think after being told that women typically perform worse than men on the task you just completed?" They had a short answer blank in which to record their answer.

Procedure

The study was completed in person, but all of the study was presented through a computer. Qualtrics Survey Software was used to obtain informed consent and implement the study. Participants were randomly assigned to either a control condition where they would not receive a threat or the experimental condition where they would be presented with a stereotype threat.

Those assigned to the control group began the study (1) did the Trait portion of the STAI; (2) watched the first instructional video and completed a practice trial; (3) watched the second instructional video, which indicated they would have 60 seconds rather than 30 seconds to study and memorize the array of objects; then, (4) did the first of the three OLM task trial, which was considered to be the baseline trial; (5) watched the third video, which reiterated the initial

instructions. (6) Participants were then given the attention check questions; after responding to these questions, (6) a second OLM task trial was administered—test trial 1, followed by (8) the third OLM task trial—test trial 2. (9) The State portion of the STAI was then administered, followed by the GIQ. (10) Testing concluded with the demographic questions.

Participants assigned to the experimental group experienced the same procedure. However, in step 5, where they watch the third video, instead of reiterating the instructions in video 2, they were presented with the stereotype threat. Participants in the experimental condition also had an 11^{th} step, where they were asked to briefly state their thoughts about the presentation of the stereotype threat, before completing the demographic questions.

Design

Experimental research methodology was used in this project. A one-way between subjects design was implemented. The independent variable was the presentation of the stereotype threat: one group received the threat (experimental) and one group did not receive the threat (control). The dependent variables were the performance on the OLM task, the adapted STAI score, and GIQ score. Demographic information was also obtained.

Performance on the OLM task was measured in three ways: non-error sum, total hits, and total false alarms. Performance on test trials 1 and 2 were combined. Total hits (H) was the sum of the number of objects correctly identified as having moved in test trials 1 and 2; scores could range from 0-21. Total false alarms (FA) was the sum of the number of objects incorrectly identified as having moved; scores could range from 0-21. The no-error sum was the number of hits (objects correctly identified as having moved), from test trials 1 and 2, minus the number of false alarms (objects incorrect identified as having moved); scores could range from -21 to +21.

The baseline trial, or trial 1, performance was measured as a no-error score by subtracting the number of false alarms from the number of hits on that trial; scores could range from -21 to +21.

Manipulation check

As both a manipulation check and attention check, participants were asked a series of questions regarding the two sets of instructions they heard. Women in the threat condition had an additional question asking what gender performed better on the spatial task they were about to complete after they had been presented with the threat. Those who indicated women were better at the task were excluded from the final data analyses.

Additionally, the baseline scores were used to ensure that there were no significant differences in the participant's ability to do well on the task. An independent t-test was conducted on the OLM task non-error performance from the baseline trial. At an alpha level of.05, there was no significant difference between the groups pre-threat presentation, $t(97) = .348, p = .73, d = .069$; women in the threat condition did not perform significantly worse (M=7.80, SD=5.06) than women in the no-threat condition (M=8.14, SD=4.79). Independent means t-tests were run on the number of hits and false alarms as well. At an alpha level of .05, there was no significant difference in hits between the threat group pre-threat presentation, (M= 11.24, SD= 4.59) and the no-threat group (M=11.4, SD=4.40), $t(97) = -.17, p = .864, d = .03$. There was also no significant difference in false alarms between the threat group (M=3.45, SD=2.33) and no threat group pre-threat presentation (M=3.26, SD=3.65), $t(97) = .31, p = .76, d = .06$.

Results

Hypothesis 1—OLM Task Performance

Independent samples *t*-tests were used to compare performance between the threat and no-threat conditions on all three outcome measures: non-error sum, total hits, and total false alarms; as well as on self-reported state anxiety. An alpha level of .05 was used for all tests to determine significance. (See Table 1 for means and standard deviations of all dependent variables).

The effect of stereotype threat was not significant on the non-error sum, $t(97) = .40$, $p = .69$, $d = .080$; women in the threat condition did not perform significantly worse (M=16.45, SD=7.91) than women in the no-threat condition (M=17.12, SD=8.91).

The effect of stereotype threat was not significant on the total hits, $t(97) = .13$, $p = .90$, $d = .03$; women in the threat condition did not perform significantly worse (M=23.31, SD=7.93) than women in the no-threat condition (M=23.1, SD=7.74).

The effect of stereotype threat was not significant on the total false alarms, $t(97) = .81$, $p = .42$, $d = .16$; women in the threat condition did not perform significantly worse (M=6.86, SD=4.13) than women in the no-threat condition (M=5.98, SD=6.43).

Hypothesis 1 - State Anxiety Scores

The effect of stereotype threat was not significant on state anxiety, $t(97) = .50$, $p = .621$, $d = .10$; women in the threat condition did not show significantly lower anxiety (M = 30.89, SD = 15.42) than women in the no-threat condition (M = 32.56, SD = 17.98).

Hypothesis 2

Pearson *r* correlations were used to examine the relationship between gender identification and no-error sum, total hits, and total false hits on the OLM task, for the threat and

no-threat conditions separately. An alpha level of .05 was used for all tests to determine significance.

For the threat condition, the correlation between gender identification and no-error sum was not significant, $r(49) = .018$, $p = .90$. The correlation between gender identification and OLM task total hits was not significant, $r(49) = .019$, $p = .90$. The correlation between gender identification and OLM task total false alarms was not significant, $r(49) = .003$, $p = .98$.

For the no-threat condition, the correlation between gender identification and OLM task no-error sum performance was not significant, $r(48) = -.037$, $p = .80$. The correlation between gender identification and OLM task total hits was not significant, $r(48) = -.19$, $p = .19$. The correlation between gender identification and OLM task total false alarms was not significant, $r(48) = -.18$, $p = .22$.

Hypothesis 3

Pearson *r* correlations were used to examine the relationship between state anxiety and OLM task non-error sum, total hits, and total false alarms for both the threat and no-threat conditions. An alpha level of .05 was used for all tests to determine significance.

For the threat condition, the correlation between state anxiety and OLM task no-error sum was not significant, $r(47) = -.13$, $p = .38$. The correlation between state anxiety and OLM task total hits was not significant, $r(47) = .11$, $p = .47$. The correlation between state anxiety and OLM task total false alarms was not significant, $r(47) = -.046$, $p = .77$.

For the no-threat condition, the correlation between state anxiety and OLM task no-error sum was not significant, $r(48) = -.034$, $p = .82$. The correlation between state anxiety and OLM task no-error sum was not significant, $r(48) = .057$, $p = .67$. The correlation between state anxiety and OLM task no-error sum, was not significant $r(48) = .115$, $p = .43$.

Hypothesis 4

For participants in the threat condition, the correlation between state anxiety and gender identification was significant, $r(47) = -.29$, $p = .041$. The more women identified with their own gender, the less anxiety they experienced. In the no-threat condition, the correlation between gender identity and anxiety was not significant, $r(48) = -.12$, $p = .41$.

Discussion

The first hypothesis that women in the stereotype threat group would perform more poorly on the OLM task and have higher state anxiety scores was not supported. The second hypothesis that women under threat who identify more strongly with their gender would perform worse on the OLM task, was also not supported. State anxiety was not significantly correlated with performance for those under threat (i.e., the third hypothesis), although there was an expected slight negative relationship ($r = -0.13$). Finally, the fourth hypothesis that, when under stereotype threat, high levels of gender identification would be positively correlated with higher state anxiety scores was not also not supported. However, there was a mild correlation between anxiety and gender identification for women in the threat condition, but it was a negative relationship: women with higher gender identification showed lower anxiety levels when presented with a stereotype threat in the OLM task ($r = -.29$).

There was no significant difference in the performance on the OLM task of women under threat and women not under threat, regardless of how performance was measured (i.e., no-error sum, hits, or false misses). Perhaps the lack of statistical significance, and experimental power, between the threat and no-threat conditions in this study, could have been due to the presence of what has been identified as a critical mass; that is, a group of same-gender individuals (Pennington & Heim, 2016). Creating a critical mass can eliminate the effect of stereotype threat on women's mathematical performance. In Pennington and Heim's study, women were tested

individually and in groups of 3-5. Female participants who were tested individually experienced a performance decrease on the math assessment, however, those in the groups did not experience the same decrease. The researchers hypothesized that this result was due to a difference in the perception of the threat. Those tested individually saw the threat as a potential for negative evaluation in their own personal ability and the ability of their social group (gender). This additional stress of being the single representation of their social group would lead to decreased performance on an assessment. However, those tested in groups were less vulnerable to the threat because the presence of other in-group members can decrease the stress of being the sole representation of their group. Even though in the present study women took the test individually, the presence of other women in the room completing the same assessment could have been enough to elicit the protective factor of a critical mass.

The presence of a female experimenter could have also contributed a protective factor for those under threat. There has not been much research investigating the impact of the experimenter gender on test performance, however, there has been some completed regarding the race of the experimenter on a stereotype threat also involving race. Marx and Goff (2005) examined Black participants' performance on a verbal task after inducing a stereotype threat by describing the task as diagnostically accurate and asking participants to indicate their race. They found that when Black participants were instructed by a Black experimenter, they outperformed Black participants instructed by a White experimenter. The presence of the same-race instructor was a protective and motivating factor that improved their performance. Marx and Goff did not include a non-diagnostic (no-threat) condition, so it cannot be determined how the scores compare to a non-threatened participant, but the improvement in performance existed, nonetheless. If the same-race experimenter could invoke a performance increase, it's possible the

same effect exists for a same-gender experimenter, which could explain the lack of a difference between those under threat and not under threat. In the current study, it's possible that when under threat, the women were not as concerned about their performance because the potential for negative evaluation was not as great as it could have been.

There was no significant correlation between the participants' level of gender identification and their performance on the OLM task in either the threat or no-threat condition. If participants didn't see the stereotype threat as threatening to their identity as a woman due to both the presence of a critical mass and female experimenter, then the participants' performance on the OLM task would be unaffected, because both conditions had high identification with being female (no threat had an average score of 4.00 and threat had an average score of 3.99) and would benefit from the protective factors mentioned earlier. This high identification with the female gender across both conditions could help explain why no significant difference was found between the two groups.

Additionally, this study failed to find a significant difference in state anxiety among the participants in the two conditions. This could also explain why there was no significant difference found in performance. Due to the presence of the above-mentioned protective factors, participants did not feel any more anxious about their performance on the task, even when presented with the stereotype threat, meaning that their performance did not decrease as expected. Previous research has also had mixed results regarding self-reported anxiety as a mediator (Pennington, Heim, Levy, & Larkin, 2016). Researchers have had some success with self-reported anxiety on mathematical tasks and motor performance with female participants (e.g. Mrazek et al., 2011; Laurin, 2013). However, other experimenters did not find mediation indicated for self-reported anxiety on women's mathematical performance (Spencer et al., 1999;

Tempel & Neumann, 2014). According to a meta-analysis of potential mediators of stereotype threat, the use of self-report measures could make the impacts of anxiety on performance more difficult to detect, which could explain why it did not differ based on conditions in the present study (Pennington et al., 2016).

There was a mild negative correlation between state anxiety and gender identification under threat (r = -0.29), but no significant correlation in the no-threat condition (r = -0.12). Women who identified highly with their gender indicated they experienced less anxiety when presented with the stereotype threat, which is an unexpected result, although it is mild. Only 8.41% of the variance in state anxiety was accounted for by the level of gender identification for women under threat. This unexpected result could be due to the presence of a critical mass, again, and help explain why there was no impact of the stereotype threat. The group of same-gender participants provided a protective factor for the participants as they completed the task (Pennington & Heim, 2016). Women who more strongly identified with their gender could benefit more from the group presence than those with weaker identification.

The threat presented in this study was unique as it was actually in the opposite direction of what previous research had indicated, that is, it stated that men performed better than women on a task that research has shown women actually outperform men on (Voyer et al., 2007). Women tend to perform better on object location memory tasks, however, this study drew on the societal stereotype regarding male performance advantages on spatial tasks and framed this task as spatial. In the attention check, participants were asked to identify what kind of task they were about to complete and only 33 participants identified the task as involving spatial skills (16 in threat condition and 17 in no-threat condition). Other responses indicated that it was simply a memory assessment or an examination of the different performances in gender. The latter

response was only found in the threat condition. It's possible that participants didn't associate the task with spatial performance, so they didn't experience the anxiety associated with the stereotype threat based on the societal stereotype about male spatial performance.

Finally, it's possible that the next generation of women simply didn't experience this anxiety associated with a stereotype threat like previous generations. The current cultural dynamic regarding the role of women and their capabilities could have empowered the women completing the task under the threat and resulted in the lack of significant difference between the two groups. Participants in the threat group were asked to indicate their thoughts (after completion of the assessment) about being told that men typically performed better than women on the spatial task. A sample of responses seem to support this claim:

> "I think that women can do anything a man can do, maybe women can do things better than men."
>
> "Felt defensive and a need to prove the study wrong personally. I feel men and women are equals."
>
> "I think that was shocking and I am not really sure what role the gender would play, or basically how that affects how well someone does."

Perhaps this group of young women refused to accept the fact that men could outperform women simply based on gender differences and thus did not experience a stereotype threat or the associated anxiety and decreased performance at all. It is also possible that women did not care enough about this task and its real life implications and did not experience increased anxiety because of that.

There is also the possibility that the female advantage in object location memory is more well known that predicted. Although participants in the threat group indicated that men

performed better in the manipulation check, they may have only been responding to demand characteristics and not actually believed what they were told in the threat presentation.

There are also positive implications of the results of this study. This task had a female advantage, and it's possible that this advantage can't be taken away by the presence of a stereotype threat. This study does not have enough participants to say that there was no difference between the groups, but it's possible that with a larger study the power would be high enough to support that hypothesis that a stereotype threat cannot impact a skilled advantage.

Conclusion

Due to the presence of multiple protective factors such as the presence of a critical mass and same-sex experimenter, we failed to reject the null hypothesis that no difference in OLM task performance would be found on object location memory task performance between women under threat and women not under threat. However, there was a significant negative correlation between state anxiety and gender identification for those under threat. Future research should further investigate these protective factors of a critical mass and same-gender experimenter in the context of an object location memory task to identify if they do indeed decrease state anxiety typically associated with stereotype threat.

References

Boone, A. P., & Hegarty, M. (2017). Sex differences in mental rotation tasks: Not just in the mental rotation process! *Journal of Experimental Psychology: Learning Memory and Cognition, 43*, 1005–1019. https://doi.org/10.1037/xlm0000370

Casad, B. J., Hale, P., & Wachs, F. L. (2017). Stereotype threat among girls: Differences by gender identity and math education context. *Psychology of Women Quarterly, 41*, 513–529. https://doi.org/10.1177/0361684317711412

Chung, B. G., Ehrhart, M. G., Ehrhart, K. H., Hattrup, K., & Solamon, J. (2010). Stereotype threat, state anxiety, and specific self-efficacy as predictors of promotion exam performance. *Group and Organization Management, 35,* 77-107. https://doi.org/10.1177/1059601109354839

Deaux, K., Bikmen, N., Gilkes, A., Ventuneac, A., Joseph, Y., Payne, Y. A., & Steele, C. M. (2007). Becoming American: Stereotype threat effects in Afro-Caribbean immigrant groups. *Social Psychology Quarterly, 70*, 384–404. https://doi.org/10.1177/019027250707000408

Doyle, R. A., & Voyer, D. (2016). Stereotype manipulation effects on math and spatial test performance: A meta-analysis. *Learning and Individual Differences, 47*, 103–116. https://doi.org/10.1016/j.lindif.2015.12.018

Eals, M., & Silverman, I. (1994). The Hunter-Gatherer theory of spatial sex differences: Proximate factors mediating the female advantage in recall of object arrays. *Evolution and Human Behavior, 15*, 95–105.

Ecuyer-Dab, I., & Robert, M. (2007). The female advantage in object location memory according to the foraging hypothesis: A critical analysis. *Human Nature, 18*, 365–385. https://doi.org/10.1007/s12110-007-9022-0

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

Hoffman, M., Gneezy, U., & List, J. A. (2011). Nurture affects gender differences in spatial abilities. *Proceedings of the National Academy of Sciences of the United States of America, 108*, 14786–14788.

Huber, M. E., Seitchik, A. E., Brown, A. J., Sternad, D., & Harkins, S. G. (2015). The effect of stereotype threat on performance of a rhythmic motor skill. *Journal of Experimental Psychology: Human Perception and Performance, 41*, 525–541. https://doi.org/10.1037/xhp0000039

Jamieson, J. P., & Harkins, S. G. (2007). Mere effort and stereotype threat performance effects. *Journal of Personality and Social Psychology, 93,* 544-564. https://doi.org/10.1037/0022-3514.93.4.544

Kapitanoff, S., & Pandey, C. (2017). Stereotype threat, anxiety, instructor gender, and underperformance in women. *Active Learning in Higher Education, 18*, 213–229. https://doi.org/10.1177/1469787417715202

Katz, I., Epps, E. G., & Axelson, L. J. (1964). Effect upon negro digit-symbol performance of anticipated comparison with whites and with other negroes. *Journal of Abnormal and Social Psychology, 69*, 77–83. https://doi.org/10.1037/h0048496

Laurin, R. (2013). Stereotype threat and lift effects in motor task performance: The mediating role of somatic and cognitive anxiety. *Journal of Social Psychology, 153*, 687–699. https://doi.org/10.1080/00224545.2013.821098

Linn, M. C., & Petersen, A. C. (2016). Emergence and characterization of sex differences in spatial ability : A meta-analysis. *Society for Research in Child Development, 56*, 1479–1498.

Marx, D. M., & Goff, P. A. (2005). Clearing the air: The effect of experimenter race on target's test performance and subjective experience. *British Journal of Social Psychology, 44*, 645–657. https://doi.org/10.1348/014466604X17948

Mazerolle, M., Régner, I., Rigalleau, F., & Huguet, P. (2015). Stereotype Threat Alters the Subjective Experience of Memory. *Experimental Psychology, 62,* 395–402. https://doi.org/10.1027/1618-3169/a000303

McGivern, R. F., Mutter, K. L., Anderson, J., Wideman, G., Bodnar, M., & Huston, P. J. (1998). Gender differences in incidental learning and visual recognition memory: Support for sex differences in unconscious environmental awareness. *Personality and Individual Differences, 25*, 223–232.

McGlone, M. S., & Aronson, J. (2006). Stereotype threat, identity salience, and spatial reasoning. *Journal of Applied Developmental Psychology, 27*, 486–493. https://doi.org/10.1016/j.appdev.2006.06.003

Mrazek, M. D., Chin, J. M., Schmader, T., Hartson, K. A., Smallwood, J., & Schooler, J. W. (2011). Threatened to distraction: Mind-wandering as a consequence of stereotype threat. *Journal of Experimental Social Psychology, 47,* 1243–1248.

Neave, N., Hamilton, C., Hutton, L., Tildesley, N., & Pickering, A. T. (2005). Some evidence of a female advantage in object location memory using ecologically valid stimuli. *Human Nature, 16*, 146–163. https://doi.org/10.1007/s12110-005-1001-8

Pansu, P., Régner, I., Max, S., Colé, P., Nezlek, J. B., & Huguet, P. (2016). A burden for the boys: Evidence of stereotype threat in boys' reading performance. *Journal of Experimental Social Psychology, 65*, 26–30. https://doi.org/10.1016/j.jesp.2016.02.008

Pennington, C. R., & Heim, D. (2016). Creating a critical mass eliminates the effects of stereotype threat on women's mathematical performance. *The British Journal of Educational Psychology, 86*, 353–368. https://doi.org/10.1111/bjep.12110

Pennington, C. R., Heim, D., Levy, A. R., & Larkin, D. T. (2016). Twenty years of stereotype threat research: A review of psychological mediators. *PLoS ONE, 11*(1), 1-25. https://doi.org/10.1371/journal.pone.0146487

Phinney, J. S. (1992). The multigroup ethnic identity measure: a new scale for use with diverse groups. *Journal of Adolescent Research, 7*, 156-176.

Shapiro, J. R., & Williams, A. M. (2012). The Role of Stereotype Threats in Undermining Girls' and Women's Performance and Interest in STEM Fields. *Sex Roles, 66*, 175–183. https://doi.org/10.1007/s11199-011-0051-0

Silverman, I., Choi, J., & Peters, M. (2007). The Hunter-gatherer theory of sex differences in spatial abilities: Data from 40 countries. *Archives of Sexual Behavior, 36*, 261–268. https://doi.org/10.1007/s10508-006-9168-6

Silverman, I., & Eals, M. (1992). Sex differences in spatial abilities: Evolutionary theory and

data. In J. H. Barkow, L. Cosmides, & J. Tooby (Eds.), The adapted mind: Evolutionary psychology and the generation of culture, 533–549.

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

Spielberger, C. D. (1983). Manual for the State-Trait Anxiety Inventory (STAI). Palo Alto, CA: Consulting Psychologists Press.

Spiers, M. V., Sakamoto, M., Elliott, R. J., & Baumann, S. (2008). Sex differences in spatial object-location memory in a virtual grocery store. *CyberPsychology & Behavior, 11*, 471–473. https://doi.org/10.1089/cpb.2007.0058

Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology, 69*, 797–812. https://doi.org/http://dx.doi.org/10.1037/0022-3514.69.5.797

Tarampi, M. R., Heydari, N., & Hegarty, M. (2016). A tale of two types of perspective taking: Sex differences in spatial ability. *Psychological Science, 27*, 1507–1516. https://doi.org/10.1177/0956797616667459

Thompson, A. E., & Voyer, D. (2014). Sex differences in the ability to recognise non-verbal displays of emotion: A meta-analysis. *Cognition and Emotion, 28*, 1164–1195. https://doi.org/10.1080/02699931.2013.875889

Tottenham, L. S., Saucier, D., Elias, L., & Gutwin, C. (2003). Female advantage for spatial location memory in both static and dynamic environments. *Brain and Cognition, 53*, 381–383.

Voyer, D., Postma, A., Brake, B., & Imperato-McGinley, J. (2007). Gender differences in object location memory: A meta-analysis. *Psychonomic Bulletin & Review, 14*, 23–28.

Table 1

Means scores of participants in all measured variables.

	No Threat Group (Control)		Threat Group (Experimental)	
	Mean	Standard Deviation	Mean	Standard Deviation
No-Sum Error	17.12	8.91	16.45	7.91
Total Hits	23.1	7.74	23.31	7.93
Total False Alarms	5.98	6.43	6.86	4.13
Gender Identif.	4.00	.51	3.99	.47
State Anxiety	32.56	17.89	30.89	15.42

Appendix A

The adapted STAI

State: Move the slider to indicate how you felt while **completing the Spatial Memory Assessment (the task you just completed).** (Not at all---Very much so)

1. I feel calm.*
2. I feel secure.*
3. I am tense.
4. I feel strained.
5. I feel at ease.*
6. I feel comfortable.*
7. I feel nervous.
8. I am jittery.
9. I feel content.*
10. I feel steady.*

*Indicates these questions were reverse scored.

Appendix B

The adapted GIQ

Please indicate your level of agreement or disagreement with each of the following statements. (Strongly Disagree---Strongly Agree)

1. I am active in organizations or social groups that include mostly members of my own gender.
2. I have a clear sense of my gender and what it means for me.
3. I think a lot about how my life will be affected by my gender.

4. I am happy that I am a member of the gender I belong to.
5. I have a strong sense of belonging towards my own gender.
6. I feel good about my gender.
7. I have a strong attachment towards my own gender.
8. I have a lot of pride in my gender.
9. In order to learn more about my gender background, I have often talked to other people about my gender.
10. I understand pretty well what my gender membership means to me.

Appendix C

The first trial arrays for the OLM task

Figure 1. Encoding Array.

Figure 2. Test Array.