1-1-2019

The Read Scare: Reading Stereotype Threat in College Aged Males

Rachel A. Booth
Mississippi State University

The Read Scare: Reading Stereotype Threat in College Aged Males

Rachel A. Booth

Mississippi State University

The Read Scare: Reading Stereotype Threat in College Aged Males

A common topic when discussing education is the gender gap in reading abilities. The Center on Educational Policy (Chudowsky & Chudowsky, 2010) found that in each grade and in each American state, girls outperformed boys on reading tasks. When children start primary school, teachers and parents assume that boys will require additional reading assistance compared to girls (Blair & Sanford, 2004). Furthermore, this disparity continues into secondary school. The Programme for International Assessment, an international survey given to 15-year olds, reported a reading gap between the genders that rates females 38 test points ahead of males – equivalent to a year's worth of progress (OECD, 2014). This gap persists even into upper education as English majors are nearly 75% women (Siebens & Ryan, 2009).

Research suggests that this aptitude inequality could be caused by low self-concept (the belief that one can do a task well) in boys (Retelsdorf, Schwartz, & Asbrock, 2015), or by disinterest in the subject because reading is perceived as boring, difficult, or less preferable to TV (Sainsbury & Schagen, 2004). However, neither explanation attempts to address fundamental questions about the underlying reasons for boys' underperformance in reading. Rather than attributing boys' underperformance to innate gender differences in reading ability, as others have proposed, effects could be due to stereotype threat (ST).

Stereotype Threat

ST places people at risk of confirming a negative stereotype about a group with which they identify, leading to poorer task performance (Steele, 1997). The ST phenomenon occurs when there is a societal perception of a group coinciding with an awareness of that belief from those included in the group (Aronson, Fried, & Good, 2002). Common stereotypes have been well studied, such as women underperforming in math (Spencer, Steele, & Quinn, 1998), or minorities having lower intelligence (Aronson, Quinn, & Spencer, 1998).

There are large numbers of women and minorities affected by ST, resulting in performances that are not congruent with their optimal capacity in areas such as mathematics for women and intelligence tasks for minorities (Aronson, Quinn, & Spencer, 1998; Nguyen & Ryan, 2008; Steele, 1997). In their lab, Steele and Aronson (1995) found that because of ST, African Americans' standardized test scores decreased. Another probable effect of ST is a lack of diversity in science, technology, engineering, and math fields (Shapiro & Williams, 2012). Furthermore, those who suffer from ST have decreased ambition in jobs requiring the stereotyped ability (Davies, Spencer, Quinn, & Gerhardstein, 2002). They also tend to have low self-esteem (Aronson & Inzlicht, 2004) and even poor health (Blascovich, Spencer, Quinn, & Steele, 2001).

The most frequently studied targeted group is women, with research exploring how minute the ST has to be to alter performance results. Findings show ST can be induced easily – even being in a math class with a female math-confident teacher induces stress in female students who want to live up to their teacher's expectations (Shapiro, 2011). Likewise, simply having the individuals specify their own gender before a math task, having the experimenter note gender differences in the study, or the participant identifying with the task's subject matter can result in similar effects (Good & Aronson, 2008; Keller, 2007).

Though investigations of reading stereotypes in males are scarce, some work suggests that early differences in reading ability may indeed be due to stereotypes. In one of the few studies to consider reading ST for boys, Pansu et al. (2016) demonstrated a reading ST effect for boys in their study. They assigned a reading test to 80 third graders, who were either told that the purpose of the study was to evaluate the children's reading ability, or to test out a new game. The gender by condition interaction revealed that the ST effect dramatically decreased boys' reading scores. In another study, Hutchinson, Smith, and Ferris (2013) implemented ST by describing the Stroop task as a verbal task for their manipulation. They tested 187 college aged men's attentional control using error rates from the Stroop task in either a stereotyped ("this is a test measuring verbal skills of men and women'') or control condition ("this is a test measuring processing skills"), and for either mostly congruent or incongruent lists of words. The Stroop task requires participants to overcome automatically reading a color word out loud, and instead say the name of the color in which the word is printed. A congruent list of words in the Stroop task has color words like blue, red, etc. that primarily match the color ink in which the word was presented (the word blue written in blue ink), with only a few incongruent words (the word blue written in red ink). A mostly incongruent word list would have color words mostly mismatched with the ink color in which they were written (the word blue written in red ink). While both tasks require maintaining task goals, the incongruent word list does not require constant monitoring because almost every problem is incongruent. The congruent list does require individuals to constantly internally maintain their goal since only a few problems were incongruent. This mean they do not have any external reminders about their task goal, nor do they have enough practice to become accustomed to the word and ink color conflict. Stereotyped males with low working memory capacity (WMC) made significantly more errors on the Stroop task than their high WMC counterparts, but only on the mostly congruent list. This condition, problem type, and WMC interaction demonstrates ST effects for the task requiring more attentional control and no ST effects for the task requiring less attentional control. It is important to note that the study found significant differences between control and stereotyped males with a verbal skills prompt but that it used an attentional control task. First, this study shows related evidence for college-aged males suffering from reading stereotype threat. Second, the Stroop task measures attention, not verbal ability, and so it is still unknown if reading performance may suffer for males under ST.

While there has been little consideration for ST in adult males, research dedicated to identifying the mechanisms underlying males' underperformance at reading is crucial, particularly when considering the necessity of reading in daily life. The present study will therefore examine whether males believe in the negative male stereotypes for reading tasks, and whether these negative stereotypes affect reading ability.

In each of the aforementioned groups, subjects only underperformed when under ST. The abundance and range of ST research lends confidence to the theory that ST interferes with task performance.

Working Memory Capacity

In well-studied domains like mathematical performance in women, the existence and activation of ST has been established thoroughly enough that the mechanisms underlying ST have become a focus of research. In particular, studies have explored the role of WMC, or the ability to simultaneously maintain and process information, in ST. Beilock, Rydell, and McConnell (2007) demonstrated that WMC may be a primary mechanism underlying the effects of ST. Participants in both a control and ST condition were presented with math problems which demanded high or low WMC to solve. There was an interaction between ST and problem-demand, with high WMC demand problem performance decreasing under ST, demonstrating that ST took a toll on WMC. Their results indicated that WMC resources are depleted by ruminations over the perceived stereotype and that depletion in WMC leads to a decrease in women's math solving abilities.

Hutchinson, Smith, and Ferris (2013), when testing ST and attentional control with the Stroop task, also measured WMC. The study found that stereotyped participants performed worse on the Stroop task when they had low WMC. They attributed their results to the distraction hypothesis. This viewpoint maintains that low WMC individuals are distracted and thus are susceptible to ST. In the frame of Kane and Engle's (2003) dual-process theory, these individuals are unable to give attention to maintaining the internal goals needed to successfully complete a task.

Counterintuitively, it may actually be high working memory participants who are most affected by ST. Work examining choking under pressure suggests that high WMC subjects ordinarily use their capacity to employ more complex strategies and effective strategies on mathematical tasks (Beilock et al., 2007; DeCaro & Beilock 2010). Under high-pressure situations, however, some of this capacity is dedicated towards rumination, making these complex strategies more difficult to carry out. Higher WMC individuals are thus reduced to using low WMC strategies, removing their previous advantage. Low WMC individuals, who already tend towards simpler strategies, are less impacted. In other words, high pressure makes high WMC participants perform similarly to low WMC participants. If ST obeys similar mechanisms to choking under pressure, it may interact with WMC in the same way and strategy use will be rendered irrelevant under ST.

Because WMC is already known to be linked with reading ability, it is reasonable to expect similar ST effects on reading problems. Past research has found that scores on complex memory tasks, in both children and adults, predict reading achievement (Cain, Oakhill, & Bryant, 2004; Swanson, 2003; Swanson & Howell, 2001; Waters & Caplan 1996). Thus, WMC should predict reading scores under normal circumstances. However, under ST, there are two possible outcomes. In light of the Hutchinson, Smith, and Ferris (2013) study, it may be low WMC individuals who perform worse on the reading task because they may be more distracted by the stereotype. Oppositely, it could be high WMC individuals who react to the stereotype in a way that eliminates their high WMC strategies. If the latter were to happen, the correlation between WMC and reading scores will not exist. Stereotyped high WMC individuals will perform like low WMC individuals, and low WMC will continue to perform like control low WMC individuals. In sum, WMC would act as a moderator through which the ST interferes with strategy use in high WMC males.

The Present Research

The gender reading gap is reminiscent of prior misconceptions of female underperformance on mathematical tasks. In that scenario, a body of literature has demonstrated that many gender differences in math ability are caused by negative stereotypes towards females (Good & Aronson, 2008; Keller, 2007; Nguyen & Ryan, 2008; Shapiro, 2011; Spencer, Steel & Quinn 1998).

The present study sought to determine if ST impacts reading task performance for college males. Male college students completed a reading comprehension task, either with or without instructions intended to induce ST. They additionally completed two WMC tasks. If ST is truly the cause of decreased reading scores among males, then high WMC men under ST should score lower on the reading task than high WMC men in the control group, with low WMC men unaffected by the manipulation. In contrast, neither control nor experimental female reading scores should be affected.

Method

Design

The study was a 2 (gender of participant) X 2 (ST or no ST) between-subjects factorial design.

Participants

Two hundred twenty-nine students (108 female, M =18.84 years old) participated in this study for course credit. Participants were randomly selected for the control (n =102) or stereotyped (n =126) groups. Within those groups there were 54 control males and 66 stereotyped males. All participants were native English speakers. In order to run the correlation analysis planned, 30 participants were collected in each subgroup for a total of 120 participants. Due to WMC task noncompletion relating to time constraints and technical difficulties, 62 participants were excluded from analysis, and due to suspecting the manipulation, 47 participants were excluded from analysis. Therefore, the authors collected data from an additional 109 participants. From the 120 participants included in analysis, 13 did not fully complete both the WMC tasks but did complete at least one, giving enough data to compose a z-score for their WMC composite score.

Materials

The Gates-MacGintie Reading Test (GMRT). The GMRT (Gates & MacGintie, 1993) is a commonly used reading comprehension task. The GMRT is comprised of a reading comprehension task and a vocabulary task. Both tasks were used in this experiment. In the first portion, participants read 11 passages and then were asked to answer between 3 to 5 questions for each passage. Gates and MacGinitie created the passages from Arts and Science texts. The reading comprehension questions are designed to test both text-based and inferential comprehension. Immediately following the reading comprehension portion, participants were asked to read 45 short phrases, all of which contained one underlined word. They were then given a list of five options and asked to choose which of the options best served as a synonym for the previously underlined word. Form S, level AR was used.

The Operation span. The automated Operation span (Ospan; Unsworth, Heitz, Schrock, &; Engle, 2005) is a commonly used test of WMC. The automated Ospan requires participants to solve a series of math problems interleaved with unrelated letters which they are expected to remember. Following the math/letter presentation, they are asked to recall the letters in the correct serial order. Sets range from three to seven letters in length, with three sets of each length. The computer scored participants' accuracy by summing all letters recalled in the correct serial order. All subjects received the exact same Ospan sets but in a randomized order.

The Symmetry span. The automated Symmetry span (Sspan; Unsworth, Redick, Heitz, Broadway, & Engle, 2009) is another common WMC test. In this computer-automated task, subjects decide if varying shapes are vertically symmetrical. In between images of shapes to judge, red squares pop up on a random location in a grid and must be remembered. The red squares are recalled after participants judge several shapes and see several red squares, by clicking on the appropriate spaces in a grid in the correct serial order. Each of these lists include two to five squares, with three trials of each length. The Sspan consists of three trials.

The computer scored participants' accuracy by measuring the number of correct cells subjects clicked in correct order. Scores were calculated by summing the number of red square locations correctly recalled in the correct order.

Demographics Sheet. Following the three tasks, participants received a demographics sheet to collect data on gender, major, and age. Further, thirteen questions regarding academic reading exposure, reading identification, and awareness of the stereotype accompanied the general demographic questions.

Manipulation check. The demographic sheet included three items: “On a scale from 1-10, how much do you agree with this statement: ‘The ability in reading comprehension of my gender is worse than the ability of the other gender.’”(Martiny, Roth, Steffens, & Croizet, 2012); “On a scale from 1-10, how much do you agree with this statement: ‘I was motivated to perform well on the reading comprehension task to help show that I am good at reading.”’(Yeung, von Hippel, 2008); and “On a scale from 1-10, how much do you agree with this statement: ‘I was motivated to perform well on the reading comprehension task to help show that my gender is good at reading.”’

Each of the three ratings were able to serve as a manipulation check.

Procedure

The study was administered at Mississippi State University. Participants were recruited through the SONA-systems website and were given course credit as compensation.

Sessions held between one and five subjects at a time. Instructions for both conditions were simultaneously presented both verbally by the experimenter and through a physical set of instructions on a sheet of paper. ST instructions for the reading task were adapted from Beilock et al. (2007). Both groups read and heard the following:

> “We are interested in reading comprehension. Most college subjects require advanced reading skills. However, not much is known regarding the mechanisms of reading abilities. This study is meant to discover what makes some people better at reading than others.”

In the control condition, participants also heard:

> “Your performance on the reading problems you are doing today will be compared to other students in the nation.”

In the experimental condition, subjects also read and heard:

> "As you may know, at most schools female students outnumber male students in majors requiring strong reading skills. There also seems to be a growing gap in academic performance between the two sexes. Unfortunately, there has not been a good explanation for this. Your performance on the reading problems you are doing today will be compared to female responses across the nation. Our specific research question is whether females are superior at all types of reading problems or only certain types."

Participants first completed the GMRT. Then, each participant completed both the Ospan and the Sspan, with the order of WMC task presentation counterbalanced. After completing the post-test questionnaire, participants were debriefed.

Results

A total of 120 participants completed all tasks and did not suspect the manipulation.

On the questionnaire participants were asked if they believed the manipulation, yes or no. Table 1 shows the distribution of participants who believed that a reading stereotype existed for males versus those who did not. Participants' responses demonstrate that, though not a pervasive belief, there is a belief that males are poorer readers than females. Therefore, two sets of analyses were performed – one on all participants, and one on only those participants who believed in the stereotype.

Table 1

Belief vs. Non-belief in Stereotype

Condition	Belief	Non-Belief
Control	38	25
Stereotype	31	26
Total	69	51

Table 2 shows the average reported ratings for the manipulation check questions. In the first question, males and females indicate that they disagree that their gender is worse at reading than the other gender. Though male participants did not think their gender's ability was poorer at reading than females' ability, the second and third questions reveal they did feel personally motivated to show they were good at reading, and motivated to show their gender was in the second question. In the second question, both genders in both conditions indicated they felt motivated to show they personally were good readers. In the third question, both genders in the ST condition indicated they felt motivated to perform well to show their gender was good at reading. Table 2 signifies that all participants felt motivated to do well on the tasks.

Table 2

Average Manipulation Check Reports

Question	Condition	Average Male Score	Average Female Score
The ability in reading comprehension of my gender is worse than the ability of the other gender.	Control	4.43	1.96
	ST	4.70	2.38
I was motivated to perform well on the reading comprehension task to help show that I am good at reading.	Control	7.24	7.62
	ST	6.86	7.59

I was motivated to perform well on the reading comprehension task to help show that my gender is good at reading.	Control	3.03	5.27
	ST	6.21	6.31

Table 3 demonstrates the correlation among the four tasks given to participants. Due to the high correlation between GMRT comprehension and vocabulary scores, those scores were combined to give a single composite GMRT score. Both the Ospan and the Sspan were given to measure WMC, and scores for those tasks were also combined to create a single composite score for WMC to remove the task specific variance.

Table 3[1]

Correlations Among Tasks

	1	2	3	4
1. GMRTComp	1	.665**	.250**	.205*
2. GMRTVocab	.665**	1	.191**	.315**
3. Sspan	.250**	.191*	1	.479**
4. Ospan	.205*	.315**	.479**	1

**. Correlation is significant at the .001 level (2-tailed).

*. Correlation is significant at the .05 level (2-tailed).

Table 4 includes the mean z-score WMC and GMRT for each group of participants.

[1] Tasks were also analyzed independently but showed the same pattern of results as the composite scores, so the composite scores were used.

Table 4

Descriptive Statistics

Task	N	Condition	Gender	Mean	Std. Deviation
SspanTotal	34	Control	Male	.39	.78
OspanTotal	33	Control	Male	.21	.90
CompAccuracy	37	Control	Male	.37	.65
VocabAccuracy	37	Control	Male	.50	.73
SspanTotal	27	ST	Male	-.01	1.01
OspanTotal	22	ST	Male	.17	1.23
CompAccuracy	28	ST	Male	.39	.67
VocabAccuracy	28	ST	Male	.39	.74
SspanTotal	23	Control	Female	.21	.56
OspanTotal	24	Control	Female	.11	.81
CompAccuracy	26	Control	Female	-.01	.92
VocabAccuracy	26	Control	Female	-.19	.78
SspanTotal	26	ST	Female	-.32	1.1
OspanTotal	28	ST	Female	-.09	.89
CompAccuracy	29	ST	Female	.21	1.02
VocabAccuracy	29	ST	Female	.10	.91

The raw scores for WMC and GMRT were converted to z-scores, then a composite was made by averaging the *z*-scores. The composite scores were used for analysis. For the WMC tasks, if a participant was missing a value for one of the two tasks, the score on the other task was inserted as the second value. Therefore, though less than 120 participants completed both WMC tasks, 120 participants had two scores to receive a composite WMC task score. A general linear model with gender and condition as fixed variables, working memory as a covariate, and GMRT composite score as an outcome variable demonstrated a main effect of gender, with male participants outperforming females on the GMRT, $F(1, 119) = 4.81$, $p = .03$. This result indicates that the males in the study showed better reading skills than the females (Figure 1). However, there was no main effect of condition on performance, $F(1, 119) = 2.96$, $p = .09$, meaning ST did not impact reading comprehension performance (Figure 1). Additionally, gender and condition did not interact, $F(1, 119) = 1.06$, $p = .31$, indicating neither of the genders were more affected by the stereotype than the other. Working memory did predict overall GMRT performance, $F(1, 119) = 12.70$, $p < .001$, as is in line with past research linking WMC and reading abilities, but WMC did not interact with gender or condition, $F(1, 119) = .27$, $p = .60$, and $F(1,119) = .05$, $p = .82$, respectively. Nor was there a three-way interaction among gender, condition, and WMC, $F(1, 119) = .69$, $p = .41$, demonstrating that gender and condition did not impact WMC's ability to predict GMRT performance (Figure 1).

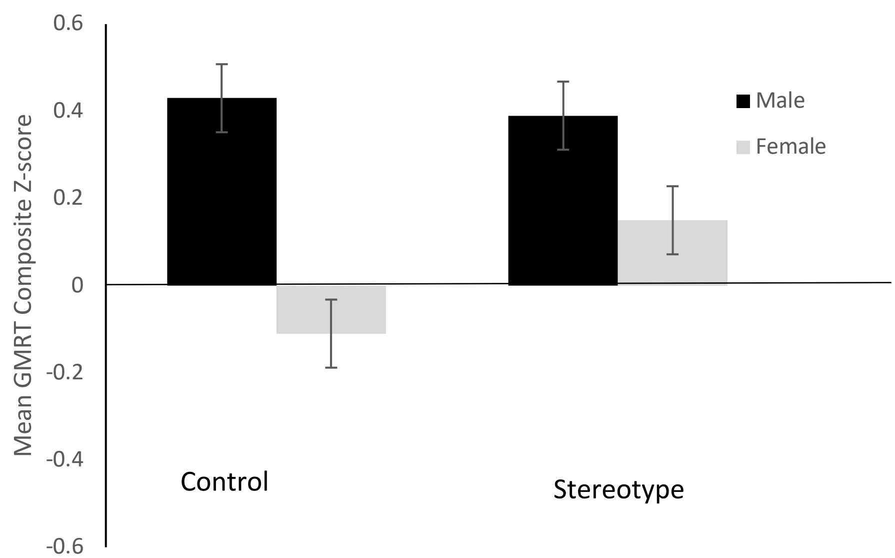

Figure 1. GMRT mean scores of all participants. Standard errors are represented in the figure by the error bars attached to each column.

The above analyses were performed again with the subset of participants who believed the stereotype. A general linear model with gender and condition as fixed variables, working memory as a covariate, and GMRT composite score as an outcome variable demonstrated a main effect of gender, with male participants outperforming females on the GMRT, $F(1, 50) = 6.36$, $p = .02$ (Figure 2). There was still no effect of condition on performance, $F(1, 50) = 1.87$, $p = .18$, therefore even among those who believed the stereotype, the stereotype did not affect their performance (Figure 2).

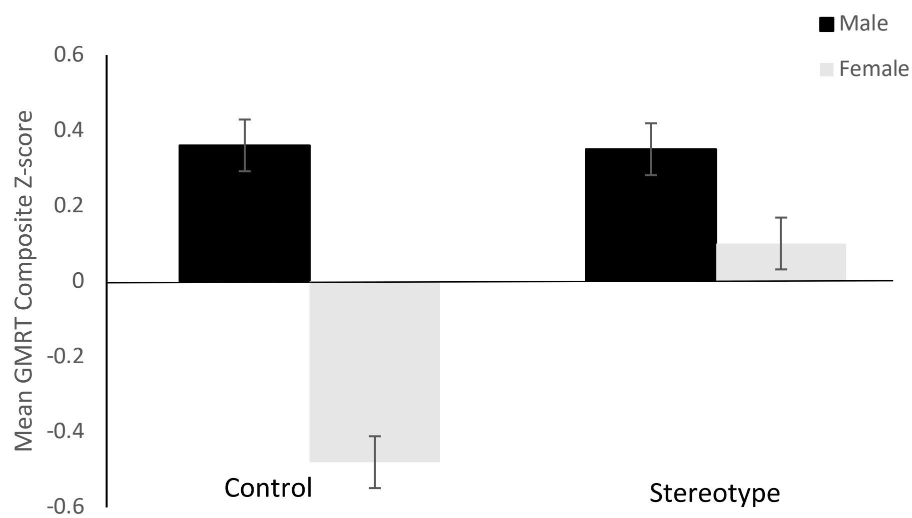

Figure 2. GMRT mean scores of belief participants. Standard errors are represented in the figure by the error bars attached to each column.

The conditions did not interact with gender, $F(1, 50) = 1.33$, $p = .26$. There was a main effect with WMC and the GMRT where WMC predicted GMRT performance, $F(1, 50) = 1.72$, $p = .02$. Unlike with the prior analyses, in this belief subset WMC and gender interacted, $F(1, 50) = 5.58$, $p =.02$. Since WMC and gender interacted in the participants who believed, a follow- up correlation showed that for males WMC and reading performance was positively correlated, $r(65) = .46$, $p < .001$. For females, WMC and reading performance was not correlated, r(55) = .26, $p = .07$. Once again, WMC and condition did not interact, $F(1, 50) = .85$, $p = .35$, nor was there an interaction among WMC, gender, and condition with those who believed the stereotype. F(1, 50) = .88, $p = .35$.

A follow-up *t*-test was conducted for all participants. There were no differences in WMC between male and female participants. (Figure 3), $t(1, 119) = 1.62$, $p < .60$.

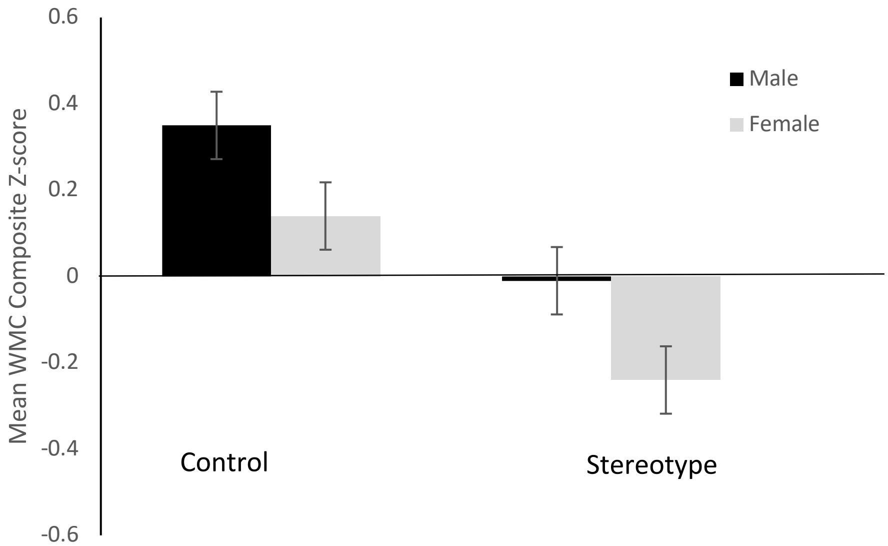

Figure 3. WMC mean scores of all participants. Standard errors are represented in the figure by the error bars attached to each column.

In the sub-section pf participants who believed the stereotype, a follow-up *t*-test with WMC as an outcome variable and gender as the independent variable was conducted. The *t*-test demonstrated females' WMC scores were significantly worse than males among those who believed the stereotype, $t(1, 49) = 4.39$, $p = 0.04$ (Figure 4).

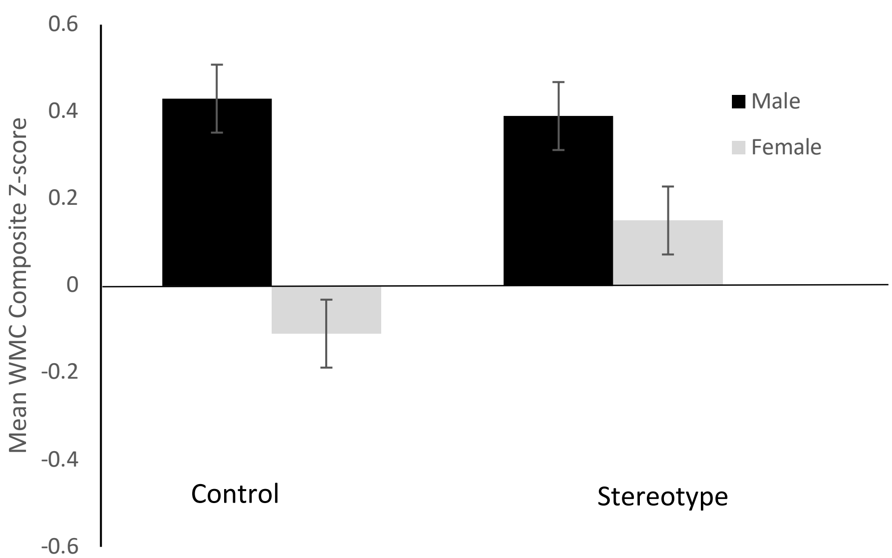

Figure 4. WMC mean scores of belief participants. Standard errors are represented in the figure by the error bars attached to each column.

The GMRT tests for both inferential and text-based questions and so a post-hoc analysis was conducted to determine if condition or gender would predict performance on each question type. Text-based questions are questions where a participant must recall a fact which was explicitly stated in the passage while inference-based questions are questions where a participant must surmise the meaning based off the text, but without the answer being in the passage. Coding was based off of prior, unpublished work by Tricia Guerrero (Guerrero, Personal Communication, July 26, 2018) and proceeded to use her scheme. If the answer could be found verbatim in the text, the question was labeled as text-based, while if the answer was not found verbatim in the text, it was labeled as inference-based. A general linear model of inference-based questions and text-based questions found no main effects with condition for inference based questions $F(1, 115) = .26$, $p = .61$, or for text-based questions $F(1,115) = 1.45$, $p = .23$. Nor was there a main effect for inference and gender $F(1, 115) = 2.83$, $p = .09$. However there was a main effect of text-based questions and gender $F(1, 115) = 5.52$, $p = .02$. A follow up *t*-test showed males outperformed females on text-based questions, $t(1, 114) = 11.49$, $p =.03$.

Discussion

The current study considered if ST existed for college-aged males on reading tasks and if WMC would moderate the relationship between condition and performance. The hypotheses for this study predicted that a stereotype did exist and ST would negatively impact performance regardless of high WMC individuals. However, those hypotheses were not confirmed. Although participants generally believed in a reading stereotype, this belief was not universal. A working memory by condition interaction demonstrated that for both control and stereotyped participants, working memory significantly predicted performance on the GMRT. Surprisingly, neither men nor women reacted to a negative male stereotype in this study.

There are many possible explanations for why Hutchinson, Smith, and Ferris (2013) and Pansu et al. (2016) found ST effects and the current study did not. One such potential reason is the role of expertise and automatized tasks. Hutchinson, Smith, and Ferris used the Stoop task in their study. Participants must report which color they are seeing and not reading, which requires participants to go against their automatized functions to perform well. Similarly, Beilock et al. (2007) used modular arithmetic tasks which are novel to most college students. These problems require students to solve subtraction and division problems then make a truth statement about the remainder. Although the math is simple arithmetic the set-up of the problems and the truth statements demand participants' focus. Both the Stroop and modular arithmetic tasks are difficult and require the participants' attention. In contrast, the current study used paragraph reading samples. For college students, reading paragraph passages could be an automatized task due to the reading expertise required of them in their college coursework. If so, then the GMRT's paragraph readings would not challenge their cognitive processes such as WMC.

If the null results are in fact due to reading expertise, then that would fit the current body of literature on ST. Since ST impacts performance by using WMC through rumination (Beilock et al. 2007), if the task itself does not challenge participants' WMC then the ruminations caused by ST will not take away any task-related resources. This difference in ST effects on level of difficulty was demonstrated in Hutchinson, Smith, and Ferris (2013) where only in the attention demanding congruent task did participants react to the ST. In their less attentionally taxing, incongruent trials, ST had no impact on participants. The math tasks in Beilock et al. (2007) were strenuous as well. This would also explain why the present study's finding do not match with Pansu et al. (2016) since reading expertise cannot be achieved with their participants of third graders, but it is reasonable to expect it from college aged adults.

However, the explanation of the GMRT task and ST not competing for WMC resources does not initially explain why WMC still correlated with reading performance but does make sense if males were not under ST. Males indicated that they believed the stereotype, but they may have not felt ST while completing the tasks. Males may have felt that while males as a gender were worse readers, they personally were not worse readers. Therefore, rather than increase task difficulty, changing the stereotype instructions may show ST effects. A Likert scale assessing if males considered their performance equal to their female counterparts' performance may be given at the end of task completion. Should males indicate they do not feel personally worse at reading, future studies may add to their ST instructions to include that even high performing males cannot compare to low performing females. Further, performance percentages may be shown on participants computer screens, with the ST condition males shown they are performing poorly. If males were not under ST because they did not feel personally worse, and future studies assert that these males are not performing as well as their female peers, ST effects may appear.

This study did contain limitations. The present study did not include a pre-post test for WMC. Since ST was introduced before any measurement of WMC, it is not evident whether the relatively small number of participants who correctly completed the WMC tasks was due to ST effects on WMC or the population's WMC abilities. Additionally, the GMRT is a short reading task which may not have been challenging enough to compete with ST for WMC resources.

To further examine the null results, future studies may incorporate lengthier, more difficult texts to explore the possible problem of participants' expertise within this study. It is possible that an effect may be found, but only with texts which require attentional control. Additionally, science texts may be used to examine an interaction between males and reading, and women and science. In this instance, males may negatively react to a stereotype concerning reading and positively react to a stereotype concerning science, and vice versa for females.

Though the current study did not demonstrate ST for college aged males in reading, the study's limitations must be considered. Males may still suffer from ST in reading but only when the reading texts demand their attentional control, or when males feel personally stereotyped. Understanding how males are able to cognitively combat ST may help mitigate consequences for those who are affected by ST. It may be the result of being in a privileged environment or ST may trigger a different response within males, such as competitiveness rather than self-doubt. If researchers can determine whether it is sociological factors or cognitive features that protects individuals against ST effects, then future literature may tackle how other stereotyped people may also use that protection.

The study's hypotheses were not confirmed and in finding unexpected results, questions as to the nature of ST arise. Women and minorities negatively react to ST when their social groups are stereotyped and males in this study did not. Therefore, researchers may determine if males require a personal stereotype to have ST effects, or if males are fundamentally less susceptible to ST. These are multi-faceted questions, pulling from sociological perspectives and individuals' differing upbringings and environments. This study brings into focus the complex and malleable nature of stereotypes.

Role of the Funding Source

This project was funded by Mississippi State University Shackouls College. After funding the project, they did not play a role in the execution of the study.

References

Aronson, J., Quinn, D. M., & Spencer, S. J. (1998). Stereotype threat and the academic underperformance of minorities and women. *Prejudice: The Target's Perspective*, 83-103.

Aronson, J., & Inzlicht, M. (2004). The ups and downs of attributional ambiguity: Stereotype vulnerability and the academic self-knowledge of African American college students. *Psychological science*, *15*(12), 829-836.

Aronson, J., Fried, C. B., & Good, C. (2002). Reducing the effects of stereotype threat on African American college students by shaping theories of intelligence. *Journal of Experimental Social Psychology*, *38*(2), 113-125.

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.

Beilock, S. L., Rydell, R. J., & McConnell, A. R. (2007). Stereotype threat and working memory: mechanisms, alleviation, and spillover. *Journal of Experimental Psychology: General*, *136*(2), 256-276.

Blair, H. A., & Sanford, K. (2004). Morphing literacy: Boys reshaping their school-based literacy practices. *Language Arts*, *81*(6), 452-460.

Blascovich, J., Spencer, S. J., Quinn, D., & Steele, C. (2001). African Americans and high blood pressure: The role of stereotype threat. *Psychological science*, *12*(3), 225-229.

DeCaro, M. S., & Beilock, S. L. (2010). The benefits and perils of attentional control. In M. Csikszentmihalyi & B. Bruya (Eds.), *Effortless attention: A new perspective in the cognitive science of attention and action* (pp. 51–73). Cambridge, MA: MIT Press.

Gates, A. I., & MacGinitie, W. H. (1971). *Gates-MacGinitie reading tests*. Teachers College Press, Columbia University.

Cain, K., Oakhill, J., & Bryant, P. (2004). Children's reading comprehension ability: Concurrent prediction by working memory, verbal ability, and component skills. *Journal of educational psychology*, *96*(1), 31.

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.

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.

Hutchison, K. A., Smith, J. L., & Ferris, A. (2013). Goals can be threatened to extinction: Using the stroop task to clarify working memory depletion under stereotype threat. *Social Psychological and Personality Science*, *4*(1), 74-81.

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.

Keller, J. (2007). Stereotype threat in classroom settings: The interactive effect of domain identification, task difficulty and stereotype threat on female students' maths performance. *British journal of Educational Psychology*, *77*(2), 323-338.

Lefavrais, P. (1986). La pipe et le rat ou l'évaluation du savoir-lire du Cours Préparatoire à l'Enseignement Supérieur. *Paris: EAP*.

Levy, B. (1996). Improving memory in old age through implicit self-stereotyping. *Journal of Personality and Social Psychology*, *71*(6), 1092.

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.

Martiny, S. E., Roth, J., Jelenec, P., Steffens, M. C., & Croizet, J. C. (2012). When a new group identity does harm on the spot: Stereotype threat in newly created groups. *European Journal of Social Psychology*, *42*(1), 65-71.

Nelson, M. J., Brown, J. I., & Denny, M. J. (1960). *The Nelson-Denny Reading Test: Vocabulary, Comprehension, Rate*. Houghton Mifflin.

Nguyen, H. H. D., & Ryan, A. M. (2008). Does stereotype threat affect test performance of minorities and women? A meta-analysis of experimental evidence.

OECD (2014). (Revised ed.). PISA 2012 Results: What Students Know and Can Do Student Performance in Mathematics, Reading and Science, Volume I,. PISA, OECD. (February 2014).

Pansu, P. Regner, I., Max, S., Colé, P., Nezlek, J.B., and Huguet, P. (2016) "A Burden for the Boys: Evidence of Stereotype Threat in Boys' Reading Performance." *Journal of Experimental Social Psychology* 65 (2016): 26-30.

Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior research methods*, *40*(3), 879-891.

Sainsbury, M., & Schagen, I. (2004). Attitudes to reading at ages nine and eleven. *Journal of Research in Reading*, *27*(4), 373-386. doi:10.1111/j.1467-9817.2004.00240.

Schmader, T., & Johns, M. (2003). Converging evidence that stereotype threat reduces working memory capacity. *Journal of personality and social psychology*, *85*(3), 440.

Shapiro, J. R. (2011). Different groups, different threats: A multi-threat approach to the experience of stereotype threats. *Personality and Social Psychology Bulletin*, *37*(4), 464-480.

Shapiro, J. R., & Neuberg, S. L. (2007). From stereotype threat to stereotype threats: Implications of a multi-threat framework for causes, moderators, mediators, consequences, and interventions. *Personality and Social Psychology Review*, *11*(2), 107-130.

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*(3-4), 175-183.

Shih, M., Pittinsky, T. L., & Ambady, N. (1999). Stereotype susceptibility: Identity salience and shifts in quantitative performance. *Psychological science*, *10*(1), 80-83

Siebens, Julie, & Ryan, L., Camille (2012). Field of Bachelor's Degrees in the United States: 2009. *American Community Survey Reports.*

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

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

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.

Swanson, H. L. (2003). Age-related differences in learning disabled and skilled readers' working memory. *Journal of experimental child psychology*, *85*(1), 1-31.

Swanson, H. L., & Howell, M. (2001). Working memory, short-term memory, and speech rate as predictors of children's reading performance at different ages. *Journal of Educational Psychology*, *93*(4), 720.

Unsworth, N., Redick, T. S., Heitz, R. P., Broadway, J. M., & Engle, R. W. (2009). Complex working memory span tasks and higher-order cognition: A latent-variable analysis of the relationship between processing and storage. *Memory*, *17*(6), 635-654.

Yeung, N. C. J., & von Hippel, C. (2008). Stereotype threat increases the likelihood that female drivers in a simulator run over jaywalkers. *Accident Analysis & Prevention*, *40*(2), 667-674.