Spring 5-1-2021

Comparing the Efficacy of Two Affirmation Interventions to Reduce Stereotype Threat Effects on Women's Math Performance

Alicia Macchione

COMPARING THE EFFICACY OF TWO AFFIRMATION INTERVENTIONS TO REDUCE STEREOTYPE THREAT EFFECTS ON WOMEN'S MATH PERFORMANCE

by

Alicia L. Macchione

A Thesis
Submitted to the Graduate School,
the College of Education and Human Sciences
and the School of Psychology
at The University of Southern Mississippi
in Partial Fulfillment of the Requirements
for the Degree of Master of Arts

Approved by:

Name of Dr. Donald F. Sacco, Committee Chair
Dr. Lucas A. Keefer
Dr. Emily R. Fyfe

May 2021

ABSTRACT

Research demonstrates when math-based gender stereotypes are activated (i.e., men are better at math than women), women display comparatively poorer math performance than men, a phenomenon referred to as stereotype threat. The current thesis project evaluated two self-affirmation interventions designed to reduce the effects of stereotype threat on women's math performance. Participants completed a math test under one of four conditions: control (no explicit stereotype activation), stereotype threat (activation of gender performance stereotype) or stereotype threat combined with one of two self-affirmation manipulations. Women in the affirmation conditions either read about women's greater verbal or relational ability and were asked to write about why the trait is important to their self-concept. No omnibus effect of condition emerged, though exploratory analyses revealed several notable findings. First, we were unable to replicate stereotype threat effects; women in the stereotype threat condition performed equivalently to women in the no threat condition. Though all individual comparisons did not reach conventional statistical significance, exploratory contrasts revealed that the combined performance of women in the two affirmation conditions was greater than the combined performance of women in the two no-affirmation conditions. More specifically, the performance of women in the relational affirmation condition was greater than the combined performance of women in the other three conditions. Though performance enhancement in the affirmation conditions was consistent with study hypotheses, the relative greater benefits of relational compared to verbal affirmation ran counter to study hypotheses. No conditional effects emerged for affect, performance regret, or interest in STEM and non-STEM careers. These findings demonstrate how self-

affirmation, particularly relational affirmation, facilitates mathematics problem-solving, independent of stereotype threat activation.

ACKNOWLEDGEMENTS

I would like to acknowledge my mentors and committee members, Dr. Donald Sacco, Dr. Lucas Keefer, and Dr. Emily Fyfe, for all of their advice, feedback, and support during the development of this thesis. Without their guidance, this thesis project would not have been possible. Thank you for continuously challenging me and helping me grow as a researcher.

TABLE OF CONTENTS

CHAPTER I – INTRODUCTION

To meet employment sector needs in an increasingly technologically-driven world, U.S companies are motivated to identify and hire employees with specific training in science, technology, engineering and mathematics, referred to as STEM disciplines (Rothwell, 2013; Stewart, 2018). In response to these shifting employment trends, U.S. educational institutions at all levels are increasingly focused on training individuals for preparation to fill STEM-based employment sectors (Xie et al., 2015). Nonetheless, concerns have been raised by an insufficient supply of individuals trained to meet these STEM employment needs, driven in part by the significant underrepresentation of women in STEM fields, often referred to as the leaky pipeline problem (Dasgupta & Stout, 2014). Contributing to this problem is the fact that women are less likely to self-select into STEM careers and face more challenges when they do. One notable potential barrier to women's STEM training is a pervasive stereotype that women are simply less capable of math-based activities and when such stereotypes are salient, women tend to underperform on math tasks, referred to as stereotype threat effects, which may discourage women from persevering in STEM disciplines. To this end, the current study is designed to assess the efficacy of two self-affirmation interventions as means of reducing performance decrements for women when engaging in a math-based task.

1.1 Stereotypes and Their Effects

Stereotypes refer to relatively fixed beliefs about a particular individual or group of individuals (Cardwell, 1996), and are thus cognitive in nature. A stereotype may be predominantly positive or negative or can contain both positive and negative content (Steele, 1997; Spencer et al. 1999; Brown and Josephs, 1999; Levy, 1996; Shih et al.

1999). In response to a complex social world, humans categorize social agents as a means of simplifying their social environment to allow for efficient processing of others; as such, stereotypes are at their core organizational. Individuals or groups are typically stereotyped based on their sex, socioeconomic status, race and ethnicity, age, gender identity, sexuality, nationality, etc. Stereotyping is essentially inferring that one individual has a set of characteristics and abilities that are shared with all other members of that particular group. For example, a common gender stereotype is that women are not as capable as men in mathematically-intensive STEM domains (e.g., Cheryan, Master, and Meltzoff 2015; Nose, Smyth, Sriram, Lindner, Devos, Ayala, et al., 2009), and underlying reasons for this belief are varied, ranging from biological differences (Benbow, Lubinski, Shea, Eftekhair-Sanjani, 2000) to social learning processes (Ceci, Williams, & Barnett, 2009).

Research has documented diverse effects that stereotypes can have over one's identity development, as well as their academic and career choices. For example, based on the existing stereotype that women are not capable of succeeding in STEM fields, women may actively choose not to pursue these fields based on the assumption that they will not succeed due to a perceived mismatch between their gender and the skills it affords and STEM disciplines and the skills they require (e.g., Master & Meltzoff, 2017). The pervasive, though largely unsupported stereotype positing gender differences in mathematical ability (e.g., Lindberg, Hyde, Peterson, Linn, 2010) can nonetheless have far-reaching influences, including creating the appearance of gender-based math ability differences not because of inherent differences in capability, but because of extra-task

emotional and cognitive processes that impede performance (e.g., Schmader, Johns, Forbes, 2008).

Stereotype threat occurs when individuals feel they are at risk of confirming a negative stereotype (Steele & Aronson, 1995). For example, when solving mathematics problems, women may fear confirming the stereotype that "women are bad at math." This risk of confirming a stereotype may increase performance regret and rumination following performance, given that their performance may negatively reflect their gender. A growing body of research has documented that stereotype threat can not only impair women's performance during mathematics problem solving (e.g., Spencer, Steele, & Quinn, 1999), but also hinder their learning of new math material (e.g., Rydell, Rydell, & Boucher, 2010). Alternatively, when the converse of a stereotype is positive, the group for whom it applies actually may benefit, or experience a performance boost. For example, when faced with a stereotype (e.g., STEM = men), men may perform better than they normally would have performed had the stereotype not been made salient (e.g., Johnson, Barnard-Brak, Saxon, & Shen, 2012). The varied effects of stereotypes are often referred to as stereotype threat and stereotype lift, and the combination of effects can further exacerbate artificial gender-based performance differences, further discouraging women from STEM pursuits.

In one of the first experimental tests of stereotype threat with women, Spencer and colleagues (1999) had men and women complete a mathematics test comprised of GRE questions. Some participants were told that the test had shown gender differences in the past, whereas other participants were told that the same test had never resulted in gender-based performance differences. When participants were led to believe that the

test produced gender differences, women scored lower compared to men. However, when participants were led to believe that the test did not produce gender differences, women scored similarly to men. Since then, many studies have replicated this finding that women underperform relative to men in threat contexts associated with math performance (e.g., Brown & Josephs, 1999; Inzlicht & Ben-Zeev, 2000; Martens, Johns, Greenberg, & Schimel, 2006; O'Brien & Crandall, 2003). Although a variety of mechanisms have been proposed, a prominent model suggests that the threatening context produces physiological, affective, and cognitive responses (e.g., arousal, emotion suppression), which in turn consume working memory resources (Schmader et al., 2008). Activities that overload working memory capacity often prevent task-relevant processing in ways that hinder learning (Sweller, Van Merrienboer, & Pass, 1998).

In an early empirical demonstration of stereotype lift, men experienced a performance boost when primed with the salient stereotype that men outperform women on mathematics problem solving (Walton & Cohen, 2003). Research has found that highlighting a negative stereotype about a specific outgroup can put members of the ingroup at an unfair advantage. More so, research has also found the salience of a stereotype that derogates an outgroup or outgroup member can increase the self-esteem of members within the ingroup (Fein & Spencer, 1997; Bandura, 1986). Research has found that when the weaknesses of an outgroup are highlighted through the presence of a stereotype, members of an ingroup may experience a performance boost based on the assumption that they should succeed at a task (Chalabaev, Stone, Sarrazin, & Croizet, 2008). By comparing oneself to the derogated outgroup, individuals may experience heightened self-esteem that in turn, may improve one's performance (Walton & Cohen,

2003). Importantly, the performance lift for the group for whom the stereotype is positive can further exacerbate the appearance of group-level capacity differences, creating an additional disadvantage for the group for whom the stereotype is negative, a particularly problematic outcome given that no innate group-level differences emerge when the stereotype is not activated. Given the negative implications of the effects of stereotype threat, there has been significant interest in identifying strategies and interventions that mitigate threat-driven performance deficits that are efficient to implement. In this paper, I will focus on one set of strategies broadly associated with self-affirmation processes that have some evidence for efficacy.

1.2. Self-Affirmation and Stereotype Threat

Self-affirmation theory suggests that individuals are highly motivated to achieve and maintain their self-image, sense of self-worth, and integrity (Steele & Liu, 1983; Steele, 1988). In other words, when one's identity is threatened, people respond by seeking defense mechanisms to restore their image, without significantly impacting their integrity (Sherman & Cohen, 2006). A self-affirmation is an act that can be used to enhance an individual's self-adequacy (Steele, 1988). For example, an act of self-affirmation could include highlighting one's strengths through positive feedback (e.g., Cohen, Aronson, & Steele, 2000). Self-affirmations can help individuals to feel less consumed with failure on performance-based assessments, and thus, may serve a critical self-protective role when the self is threatened (Koole, Smeets, Van Knippenberg, & Dijkesterhuis, 1999; Creswell, Dutcher, Klein, Harris, Levine, 2013).

In the case of stereotype threat, in which an individual's identity feels threatened, self-affirmation may serve as a buffer that can prevent one from engaging in ruminative

thinking that can interfere with performance. When a stereotype is made salient (i.e. women are bad at math), stereotyped individuals' attention may be consumed by the statement and how to respond in the face of a negative stereotype. Further, this type of ruminative thinking has been found to be detrimental in that it can unwantedly occupy one's attentional capacity and subsequently prevent goal achievement (e.g., Brunstein & Gollwitzer, 1996, Mikulincer, 1996). Ample research has found that stereotype threat places higher demands on one's mental capacity, which, in turn, hinders performance (e.g., Schmader & Johns, 2003; Rydell, Boucher, 2017). Some work has also found that the negative effects of stereotype threat can be alleviated through self-affirmation. In a standard self-affirmation study, people are often asked to write about important values/traits or are asked to complete questionnaires designed to have participants think about important values/traits (Sherman, 2013). A key aspect of the affirmation manipulation is that participants are asked to talk about a value/trait that is personally important. When introduced in a threatening context, self-affirmations can help one navigate their environment.

In an experimental test that looked at the effects of self-affirmation, Martens and colleagues (2006) had men and women complete a mathematics test. Men and women were randomly assigned to one of two conditions: stereotype threat condition or non-diagnostic test (control) condition. In the stereotype threat condition, participants completed the math test framed as a measure of one's math intelligence. However, women assigned to a third stereotype threat + self-affirmation condition were told that the math test was a measure of one's math intelligence and were also presented with an affirmation paradigm in which they were asked to affirm a valued/non-valued trait.

Women in the threat condition performed significantly worse on the math test than the women in the non-threat/control condition, while men performed similarly in the control and stereotype threat conditions. Most relevant to the current study, women in the threat/self-affirmation condition significantly outperformed women in the no-affirmation/threat condition. Thus, in other words, self-affirmation was able to eliminate the negative effect of stereotype threat on women's math performance. Further, research has also found that high power can often lead to stereotype lift effects, especially when in high-pressure situations, such as in a threat context (e.g., Galinsky, Gruenfeld, & Magee, 2003; Guinote, 2007a, 2007b; Whitson, Liljenquist, Galinsky, Magee, Gruenfeld, & Cadena, 2013; Kang, Galinsky, Kray, & Shirako, 2015). Specifically, in a study that looked at how power can impact performance in high-pressure contexts, Kang and colleagues (2015) found that when the stakes are high and psychological power is low, performance was significantly lower, mimicking stereotype threat effects. On the other hand, when the stakes are high and psychological power is also high, performance is significantly higher, mimicking stereotype lift effects. Thus, psychological power, which could come in the form of self-affirmation, has the potential to mitigate against stereotype threat effects. Specifically, self-affirmation can lead to increases in one's sense of personal power and help serve as a protective barrier for threatened individuals, such that affirming a skill unrelated to the stereotype could ultimately result in stereotype lift and further, could lead to enhanced performance compared to baseline. Critically, there is limited work regarding 1) how robustly self-affirmation is in the context of thwarting stereotype threat effects and 2) how different types of self-affirmations may be more or less effective at buffering against stereotype threat effects. The current study will serve

to address these two critical gaps in the literature on self-affirmation and stereotype threat.

CHAPTER II – CURRENT RESEARCH

The goal of the current study was to compare women's problem-solving performance under threat and non-threat contexts and attempt to mitigate against the potentially harmful effects of stereotypes on performance using two forms of self-affirmation. The proposed thesis project provides a unique expansion to existing research on stereotype threat by testing a brief intervention designed to help diminish the effects of stereotype threat on problem-solving. We tested whether it is possible to eliminate the harmful potential of stereotype threat context by immediately presenting stigmatized individuals with an affirmation opportunity that will potentially serve as a buffer against threat effects. Provided that research has found that men's math performance is not harmed in stereotype threat contexts (e.g., Marten et. al, 2006), and they are not underrepresented in mathematically-intensive, STEM fields, this study focuses exclusively on women's in math performance.

In the current study, I recruited women and assigned them to one of four conditions on a between participants basis: gender-fair (control) condition women will read a benign article; stereotype threat (ST) condition women will be primed with a statement to activate the gender-difference math stereotype; ST + Relational Affirmation (RA) women will be primed with the gendered math stereotype, read about women's superior relational aptitude and write about how this capacity is important to their identity; ST + Verbal Affirmation (VA) women will be primed with gendered math stereotype, read about women's superior verbal aptitude and write about how this capacity is important to their identity. I selected these two affirmation paradigms because past literature shows women tend to outperform men on both social intelligence and

verbal ability tasks (e.g., Korn Ferry Group [KFY] 2016; Andreano & Cahill 2009; Hyde & Linn 1988; Kimura 1992), thus these are great veridical skills to use in an affirmation paradigm with women. However, verbal ability affirmation may be more useful because it allows self-protection from stereotype threat to occur in a domain related to academic ability, which may be important when the threatened domain is also associated with another aspect of academic ability, namely math-ability. Women in all four conditions will complete a math test as the primary dependent measure. The current study will test the following predictions:

H1: Consistent with stereotype threat literature in which threat context significantly hinders performance (e.g., Doyle & Voyer, 2016), I predicted that women in the threat condition will perform significantly worse relative to women in the gender-fair (control) condition, even after accounting for prior math knowledge.

H2: Given the benefits of self-affirmation on performance (e.g., Marten et al., 2006), I predicted that women in the ST+RA and ST+VA conditions will perform better on the math test than women in the ST condition.

H3: Of the two affirmation statements, I predicted that women will perform best in the verbal affirmation (VA) condition because it provides specific affirmation associated with educational performance, rather than the broad interpersonal affirmation associated with the relational affirmation condition. By affirming oneself in an intellectual domain, the threat associated with potential underperformance in the threatened domain of math should be more effectively buffered. Though affirming relational aptitude will be beneficial, its weak association with the threatened domain will result in weaker buffering effects. It may be that the presence of verbal affirmation,

relative to relational affirmation, functions similarly to stereotype lift. Given that high power has been shown to lead to stereotype lift, especially in high-pressure situations (e.g., Kang et al., 2015), I predicted that the boosted sense of empowerment from verbal affirmation, which itself communicates elevated intellectual ability compared to relational affirmation, will translate to better performance on the math task than all three other conditions, thus resulting in stereotype lift, rather than just eliminating the performance deficit traditionally associated with stereotype threat. That is, the confidence stemming from the reminder of women's greater verbal ability will result in greater confidence in intellectual performance more generally, which will ultimately result in women demonstrating higher performance compared to baseline (e.g., Kang et al., 2015).

To test these hypotheses, women were assigned to one of the four conditions outlined above, and complete a math task. The target task was to use probabilistic reasoning to solve posterior probability problems. We selected these problems because probabilistic reasoning is critical when interpreting everyday situations, yet, many individuals struggle with this type of reasoning (e.g., Hoffrage, Kurzenhauser, & Gigerenzer, 2005).

CHAPTER III – METHODOLOGY

3.1 Participants

520 undergraduate students from a large southeastern university participated in the study for extra or partial course credit in an introductory psychology course. Data from 214 participants were excluded from analyses: 39 who provided incomplete data, 7 who selected 'male' as their gender, and 168 participants who failed the manipulation check (reflecting insensitivity to conditional assignment/stereotype threat manipulation). Of those excluded, 51 were in the no threat no affirmation (control) condition, 48 were in the threat no affirmation, 61 were in the threat relational affirmation, and 54 were in the threat with verbal affirmation condition. This resulted in a final sample for analysis comprised of 306 female participants (*M* age = 20.41 years, *SD* = 4.98 years; 209 White, 72 Black, 4 Hispanic, 3 Asian, and 20 participants who reported their race as 'Other'). A, a priori, medium effect-size power analysis using G*Power (Faul, Erdfelder, & Lang, 2007) suggests 280 participants would adequately detect effects (Cohen's F=0.20, β = 0.80), confirming the study was sufficiently powered to test hypotheses.

Why not include males?

Provided that stereotypes in math typically target females, we are interested in how self-affirmations will help female performance in a threat context. Past research has shown that men perform similarly on math tasks in threat and non-threat contexts, thus, for this study, we focused only on women's performance. Further, given that self-affirmation is not likely to impact male performance, as their identity/self-integrity is unlikely to be threatened during a math task, this study recruited females only. Additionally, provided that 1) women have historically been underrepresented from

mathematically-intensive STEM domains (not men), and 2) math stereotypes undermine women's performance which may lead them to select away from STEM, we need critical interventions for women's math performance.

3.2 Design

Women were randomly assigned to one of four conditions: gender-fair (control) condition, stereotype threat (ST) condition, stereotype threat with relational affirmation (ST+RA), or stereotype threat with verbal affirmation (ST+VA). All participants completed the same mathematics assessment. See Figure 1 for a visual schematic of the study design and procedure.

3.3 Materials

During the online session, participants completed a series of math tasks including (1) solving two baseline problems and (2) solving ten target problems. Before beginning problem-solving, participants were asked to respond to a state self-esteem scale and rate their current level of anxiety, given that performance may differ based on levels of self-esteem and anxiety. Provided that stereotype threat may increase rumination following performance and further, discourage women's pursual in future math tasks, at the end of the session, participants were asked to complete a regret assessment measuring feelings post-performance to assess performance regret which may stray women from pursuing math-based tasks in the future. Additionally, participants were asked to complete an interest assessment measuring their interest in pursuing additional problems to see if the presence of threat and/or affirmation influences future interest in math. At the very end of the session, participants filled out a demographics questionnaire and were debriefed. All the math tasks will consist of probability problems that were adapted from previous work

(e.g., Fyfe & Brown, 2018b). Problems present a data table of results (e.g., polygraph test, breathalyzer test) and ask participants to calculate the likelihood of various scenarios. These problems can be used to calculate the prevalence of a condition, as well as the predictive value of a test. Prevalence items are about the prevalence of a condition and consisted of questions such as, "How likely is it that an employee is being honest?" and "How likely is it that a driver receives a positive breathalyzer test?" Predictive items are about the predictive value of a test and consisted of questions such as, "How likely is it that an employee with a negative polygraph test is actually being honest?" and "How likely is it that a driver with a negative Breathalyzer test is actually sober?". See Appendix D for all materials.

3.3.1 Baseline Problems

The first math section contained two baseline probability problems designed to measure participants' prior knowledge. Both problems refer to a single data table and are presented in a fixed order on the same screen (see Appendix D). The first problem asks participants to calculate the prevalence of drunk drivers based on the data table (prevalence problem). The second problem asks participants to calculate the predictive value of an outcome on a breathalyzer test (predictive value problem).

3.3.2 State Self-Esteem Assessment

After solving the baseline problems, participants responded to a 16-item state self-esteem questionnaire that captured three different subscales of self-esteem: social self-esteem, performance self-esteem, and appearance self-esteem (α=.89) (see Appendix D). This measure, adapted from Heatherton and Polivy's (1991) validated state self-esteem scale, was designed to measure a participant's current level of self-esteem at a given

point in time. Self-esteem was analyzed collectively, in which composite self-esteem scores were created for each participant where higher values corresponded to higher state self-esteem. Self-esteem was also analyzed by type of self-esteem: performance, social, or appearance self-esteem. It was predicted that individuals with lower self-esteem may be especially vulnerable to negative stereotype threat effects.

3.3.3 Current Anxiety Assessment

Following the state self-esteem assessment, participants responded to a single item measuring current level of anxiety, in which participants were asked to rate the level of anxiety they experienced prior to problem-solving, using a 7-point Likert scale [*1=Not At all Anxious, 7=Extremely Anxious*]. Provided that affirmation has been shown to alleviate feelings of failure that can pressure performance, it was predicted that that affirmation will reduce anxiety that participants may experience when threatened.

3.3.4 Target Problems

The second math section contained ten probability problems, including five prevalence problems and five predictive value problems. There are five unique data tables, all of which are similar in structure to the data table on the baseline items (see Appendix D). Participants will answer a prevalence problem and a predictive value problem for each data table. However, each individual problem will be presented one at a time. The problems are presented in a fixed order. Throughout the problem-solving portion, there were also two basic "attention check" problems (i.e., What is 1/5 written as a decimal? What is 10% written as a decimal?). Incorrect answers to these questions were used as an indicator that a participant was not attempting genuine responses to the tasks, and thus, these participants' responses were excluded.

3.3.5 Regret Assessment

After the problem-solving tasks, participants responded to an 18-item regret questionnaire (α=.87) (see Appendix D). This measure was designed to assess participants' feelings about themselves and their performance on the problem-solving after completing the problems. Composite performance regret scores were created for each participant where higher values corresponded to higher performance regret. Provided that the risk of confirming a stereotype may increase performance regret and that the presence of affirmation may act serve a self-protective role (e.g., Koole et al., 1999), I predicted that when self-affirmed, especially in a threat context, participants will express lower levels of performance regret. Specifically, I predicted that the self-affirmation statements will decrease the amount of pressure that is added to the participant to succeed in order to refute the negative stereotype. The less pressure a participant experiences, the less regret they may be likely to have following their performance. It may that when participants affirm a non-math skill, they are less threatened by the potential math domain and feel less pressure to perform perfectly, thus, participants may report lower levels of performance regret following affirmation. I included the regret assessments to allow for replication of prior work's articulation of stereotype threat effects, but it may also serve as an outcome variable to test for exploratory mediation. The regret assessment will be used to test the effectiveness of self-affirmations on participants' experience of regret, based on their performance.

3.3.6 Interest Assessment

Following the regret assessment, participants responded to an 8-item interest questionnaire (see Appendix D). This measure was designed to assess participants'

interest in pursuing fields that are STEM (4 items; α=.80) versus non-STEM (4 items; α=.74). Composite STEM and non-STEM interest scores were created for each participant where higher values corresponded to higher interest. I predicted that when self-affirmed, especially in a threat context, participants will express higher levels of interest in pursuing STEM fields. Specifically, I predicted that the self-affirmation statements will decrease the likelihood of threatened individuals disengaging from the math task and feeling discouraged from pursuing math fields. The interest assessment was intended to test the effectiveness of self-affirmations on fostering participants' interest in math-related fields.

3.3.7 Demographic Questionnaire

Finally, participants completed a brief background questionnaire, containing questions about their prior experience in mathematics, their college experience (i.e., major, year, GPA), and their identity (i.e., age, gender, ethnicity).

3.3.8 Debriefing

Following the completion of the study, participants were presented with a debriefing screen, including information about the stereotype threat and affirmation manipulation (Appendix D).

3.4 Procedure

See Figure 1 for a visual schematic of the study design and procedure. Participants completed a single session online. After providing consent (Appendix D), all participants were asked to complete the state self-esteem assessment before beginning. Before beginning problem-solving, all students will read the following statement:

> "This research is aimed at better understanding what makes some people better at math than others. As you may know, there has been some controversy about whether there are gender differences in math ability. Previous research has sometimes shown gender differences and sometimes shown no gender differences. Today, you will read a relevant article from research scientists regarding their recent findings on gender and math performance. After reading, you will solve a set of ten math problems that are like the ones you just solved."

1. In the gender-fair (control) condition, participants will read a short statement from a research article that explains how recent research has found that men and women perform equally well on math tasks similar to the ones you will be completing today. After reading the short article, participants will be asked to indicate their gender. Specifically, participants will be read the following statement:

> "A recent summary of all research results finds NO consistent evidence for a gender difference in math ability. This means that men and women perform equally well on math-based assessments, such as the math section of the SAT, ACT, and GRE (e.g., Williams et al., 2020)."

After reading the statement, participants will then be told:

> "You will now complete math problems similar to the ones studied above. Your performance on these math problems will be compared to other students from across the nation. To continue to the next section, please indicate your gender below: [select female or male]."

To ensure participants were attenuating to the statement/instructions, a timer was placed that prevented participants from moving forward before a sufficient amount of time as

passed necessary to read the instructions thoroughly (~60sec). Additionally, participants were asked to answer manipulation check questions (e.g., “Research suggests who performs better on math tasks?”), to further ensure that they were properly attenuating to the manipulation. Following instructions, participants rated their current level of anxiety, and then moved on to the ten-item math assessment, followed by the regret assessment, interest assessment, and a demographic questionnaire.

2. In the stereotype threat (ST) condition, participants read a short statement from a research article that explains how recent research has found that women significantly underperform men on math tasks similar to the ones that you will be completing today. After reading the short article, participants were asked to indicate their gender. Specifically, participants read the following statement:

> “A recent summary of all research results finds consistent STRONG evidence for a gender difference in math ability. This means that men outperform women on math-based assessments, such as the math section of the SAT, ACT, and GRE (e.g., Williams et al., 2020).”

To ensure participants are attenuating to the statement/instructions, a timer was placed that prevents participants from moving forward before a sufficient amount of time as passed necessary to read the instructions thoroughly (~60sec). Additionally, participants were asked to answer manipulation check questions (e.g., “Research suggests who performs better on math tasks?”), to further ensure that they were properly attenuating to the manipulation. Following instructions, participants rated their current level of anxiety, and then moved on to the 10-item math assessment, followed by the regret assessment, interest assessment, and a demographic questionnaire.

3. In the stereotype threat + relational affirmation (ST+RA) condition, participants were given the previously mentioned threat instructions immediately followed by the relational affirmation manipulation, prior to problem-solving. Participants read the following relational affirmation passage:

> "Research has shown that women's interpersonal ability is superior to men's interpersonal ability. This means that women tend to outperform men on tasks that require interpersonal skills. For example, multiple studies have found that women consistently tend to score higher on social tasks, such as reading non-verbal expressions
>
> (e.g., Cortes and Pan, 2017, Baumeister and Sommer, 1997; Gabriel and Gardner, 1999). Further, this gender gap is often maintained over time, in that adult women have higher interpersonal ability than adult men."

After reading the relational affirmation passage, participants responded to a writing prompt asking them to describe how relational ability is important to them and how much of this ability they think they have. After completing the writing prompt, participants moved to solving the ten problems, followed by a regret assessment, interest assessment, and a demographic questionnaire. At the end of the study, participants were asked questions regarding the passage they read to confirm comprehension and attention (i.e., "According to the article, who performs better at math, men or women?).

4. In the stereotype threat + verbal affirmation (ST+VA) condition, participants were given the previously mentioned threat instructions immediately followed by the verbal affirmation manipulation, prior to problem-solving. Participants read the following verbal affirmation passage:

> "Research has shown that women's verbal ability is superior to men's verbal ability. This means that women tend to outperform men on tasks that require verbal skills. For example, multiple studies have found that women consistently tend to score higher on the verbal sections of standardized tests, such as reading comprehension, writing, and vocabulary-based assessments (e.g., Andreano and Cahill 2009; Hyde and Linn 1988; Kimura 1992). Further, this gender gap is often maintained over time, in that adult women have higher verbal ability than adult men."

After reading the verbal affirmation passage, participants responded to a writing prompt asking them to describe how verbal ability is important to them and how much of this ability they think they have. After completing the writing prompt, participants moved on to solving the ten problems, followed by a regret assessment, interest assessment, and a demographic questionnaire.

Math Task

Participants across all conditions completed the same math task. They first solved a warm-up problem, that was not scored, to acquaint them with using the multiple-choice response process. Participants then completed two baseline problems to assess their prior knowledge. This was then followed by further instructions for completing the mathematics task, which included the threat manipulation, followed by the affirmation manipulation. Following manipulations, participants moved on to solving the ten target problems, followed by a regret assessment, interest assessment, and a demographic questionnaire.

CHAPTER IV – RESULTS

Primary Analyses

I predicted that the stereotype threat activation would lead to performance deficits compared to the baseline no threat condition (**H_1**), that affirmation (regardless of type) would protect against the effects of stereotype activation (**H_2**), and of the two affirmations, verbal affirmation would be the most beneficial (**H_3**). To test my hypotheses, I first conducted a one-way ANCOVA, controlling for differences in baseline math performance, with math test performance as the dependent measure and condition (4 levels) as the independent variable. Though the omnibus analysis was not significant, $F(3, 303) = 1.65$, $p = .178$, $\eta^2_p = .016$, inspection of conditional level means revealed an interesting descriptive pattern whereby performance was greatest in the threat with relational affirmation condition and to a lesser extent the threat with verbal affirmation compared to the no affirmation conditions (stereotype threat and baseline conditions); no descriptive difference appeared to emerge between the basic stereotype threat condition and the control condition, suggesting that the current study's manipulation of threat was ineffectual (Table 1). Indeed, and contrary to **Hypothesis 1**, an independent samples *t*-test revealed that women in the threat condition ($M = 5.15$, $SD = 2.86$) performed similarly to women in the no threat/gender-fair condition ($M = 4.83$, $SD = 2.45$), $t(113) = 20.79$, $p = .85$, $d = .12$.

Based on these descriptive differences described above, I conducted two contrasts to determine if specific mean level differences emerged, independent of the non-significant omnibus analysis. In the first contrast, I simultaneously compared math performance in the two affirmation conditions to the two non-affirmation conditions. A

marginally significant effect emerged, $t(304) = 1.89$, $p = .062$, $d = .81$, indicating that participants given the opportunity to affirm performed better ($M = 5.51$, $SD = 2.72$) than those not provided the opportunity to affirm ($M = 4.99$, $SD = 2.67$). Though the relational affirmation condition ($M = 5.79$, $SD = 2.83$) did not differ significantly from the verbal affirmation condition ($M = 5.23$, $SD = 2.62$), $t(304) = 1.25$, $p = .212$, it was descriptively the highest performing condition, and significantly higher than the no threat (control) ($M = 4.83$, $SD = 2.45$), $t(304) = 2.22$, $p = .028$, and descriptively higher than in the stereotype threat condition ($M = 5.15$, $SD = 2.86$), $t(304) = 1.37$, $p = .174$. As such, I conducted a more focused contrast in which I compared the relational affirmation condition simultaneously to the other three conditions. This analysis was nearly conventionally significant, $t(304) = 1.95$, $p = .052$, suggesting that relational affirmation led to greater performance than the other conditions.[2] These results are consistent with **Hypothesis 2** but inconsistent with **Hypothesis 3**.

Secondary Analyses

Beyond the above primary analyses, I also conducted exploratory analyses for self-esteem, anxiety, regret, and interest assessments. See Table 1 for descriptive statistics for each variable and see Table 2 for correlation statistics between each variable.

Self-Esteem

For self-esteem, I expected participants with low self-esteem to be especially vulnerable to negative stereotype threat effects. However, contrary to this prediction, participants expressed above average self-esteem ($M = 3.28$ out of 5), regardless of condition, and further, condition had no effect on reported levels of self-esteem, $F(3,304) = .970$, $p = .407$, $\eta^2_p = 0.009$, suggesting that all participants expressed similar levels of

self-esteem, regardless of the presence of affirmation or threat.[1] See Table 1 for descriptive statistics.

Anxiety

For anxiety, provided that affirmation has been shown to alleviate feelings of failure that can pressure performance, I expected that affirmation would reduce anxiety that participants may experience when threatened. However, contrary to this prediction, there was no significant main effect of condition on levels of anxiety, $F(1,304) = .630$, $p = .596$, $\eta^2_p = .006$, even after accounting for pretest performance, $F(1,303) = .808$, $p = .370$, $\eta^2_p = .003$. Thus, anxiety levels did not differ based on the presence of threat or affirmation. See Table 1 for descriptive statistics.

Performance Regret

For regret, provided that the risk of confirming a stereotype may increase performance regret and that self-affirmations can act as a protective barrier (e.g., Koole et al., 1999), I predicted that when self-affirmed, especially in a threat context, participants will express lower levels of regret. Specifically, I predicted that self-affirmation will decrease the amount of pressure that is added to the participant to succeed in order to refute the negative math stereotype. Contrary to this prediction, an ANCOVA with condition as a categorical independent variable, target performance as a continuous predictor independent variable, and regret as the dependent variable, revealed there was no significant main effect of condition on levels of performance regret, $F(1,303) = .054$, $p = .983$, $\eta^2_p = .001$. However, there was a sensible main effect of actual performance on performance regret, $F(1,302) = 227.01$, $p < .001$, $\eta^2_p = .429$. To determine the direction of the relationship between actual performance and performance regret, I correlated

participants' performance with their performance regret. This revealed a sensible finding, $r(308) = -.715$, $p < .001$, such that the higher performance on the target problems, the lower performance regret that was expressed. Thus, while the presence of affirmation did not influence participants' feelings of performance regret, one's overall greater performance was sensibly correlated with the experience of less regret. See Table 1 for descriptive statistics.

Interest in STEM and non-STEM

For interest, I predicted that when self-affirmed, especially in a threat context, participants will express higher levels of interest in pursuing STEM careers and less interest in pursuing non-STEM careers, if true, would be mediated by enhanced performance in the affirmation conditions. Contrary to predictions, a MANOVA with STEM interest and non-STEM interest as separate dependent measures and condition as a between-subjects factor revealed that there was no effect of condition on either STEM interest, $F(3,304) = 0.89$, $p = .447$, $\eta^2_p = 0.009$, or non-STEM interest $F(3,304) = 1.18$, $p = .317$, $\eta^2_p = 0.012$. Thus, the presence of affirmation did not significantly affect participants' interest. See Table 1 for descriptive statistics.

CHAPTER V – DISCUSSION

This study was intended to test two self-affirmation interventions designed to reduce the effects of stereotype threat on women's math performance. Several hypotheses were not supported, while others were partially (though rather weakly) supported. Interestingly, and contrary to our predictions, we were unable to replicate basic stereotype threat effects – women performed similarly within threat and no threat conditions. Research regarding stereotype threat and women's math performance has produced mixed findings. For example, a recent study found that stereotype threat activation did not impair women's mathematical performance (i.e., Pennington, Litchfield, McLatchie, & Heim, 2018), and thus, the current study's null stereotype threat effect findings are not unprecedented. However, while we were unable to demonstrate stereotype threat effects, we did find that providing women with affirmation led to higher performance than women who were not affirmed. This finding is consistent with work showing that self-affirmation can be beneficial for boosting performance (e.g., Creswell et. al, 2013; Harris, Harris, & Miles, 2017). Provided that verbal affirmation can further validate one's intellectual abilities necessary to complete math tasks, it was predicted that verbal affirmation would be the most beneficial at boosting women's math performance, however, contrary to these expectations, affirming women in relational abilities actually resulted in the greatest performance boosts relative to women in the other conditions. It is possible that women in our sample did not inherently value verbal abilities or see them personally valuable or relevant as relational abilities, resulting in relatively weak math performance enhancement for verbal affirmation. When women in threatening contexts are asked to affirm a *valued* skill, they performed similarly to men and to women in non-

threatening contexts (e.g., Martens et al., 2006), as self-affirmation provides a protective barrier against negative threat effects. However, it could be that women value relational abilities over verbal abilities, and believe themselves to be more superior relationally, and thus, affirming these skills specifically led to the most protection from stereotype threat effects and facilitated performance. Further, given that female gender stereotypes suggest that women are more communal and relationship-focused than men (e.g., Broverman et al., 1972; Eagly and Steffen, 1984), it could be that participants found the description of women's superior in relational skills to be more plausible than superior verbal skills, making relational affirmation more influential on performance than verbal affirmation.

Secondary predictions including anxiety, interest, and regret assessments produced no condition level effects, suggesting that the presence of threat and/or affirmation did not significantly influence these factors. Given that the stereotype manipulation was unsuccessful, it is not surprising that condition did not significantly influence participants' feelings of anxiety or their feelings of performance regret. However, interestingly, while the presence of affirmation, specifically relational affirmation, led to performance boosts, affirmation and performance did not influence interest in either STEM or non-STEM fields. Future research would benefit by recruiting an equal number of declared and undeclared majors, to see if those who have yet to declare might be more amenable to fluctuations in interest based on affirmation and performance. Conversely, those who have already selected a major may be less likely to show changes in interest.

5.1 Limitations and Future Directions

A primary limitation of this study was the inability to replicate stereotype threat effects; women performed similarly regardless of whether stereotypes were activated. Provided that the threat manipulation was given at the beginning of the study, and the manipulation check was not given until after problem-solving, it is difficult to know whether or not participants were actually attenuating to the threat manipulation entirely, and whether or not the activation of threat was held throughout the entirety of the study. Follow-up studies could investigate whether threat made more salient throughout the entire study (i.e., presenting threatening information on every page) could result in replication of stereotype threat effects. In addition to the potential lack of threat saliency, our demographic used (college students) could have influenced findings. It could be that college-aged participants, especially those that are within the introductory psychology pool (a STEM field) are less susceptible to the negative effects of gender-based math performance stereotypes, as they have likely already experienced overcoming barriers to reach higher education and have direct experience as a female navigating a STEM career trajectory. On that note, it could be that our demographic did not believe the information provided in the threat manipulation – although they were able to read the manipulation and confirmed that they understood what was presented by passing the manipulation check, it could be that participants did not find the information believable, and thus, were less susceptible to the threat effects. Further, research could also consider asking participants to rate the extent to which they believe the threat manipulation to understand what may have led to the failed stereotype threat manipulation.

Further, research investigating stereotype threat effects has found that effects are strongest when the domain is important to the participant's identity. In this study, we recruited psychology students, a field that while categorized as a STEM discipline is not a traditional math-intensive STEM field. Consequently, it could be that math ability is not as important to this samples' self-concept, minimizing the effect of math stereotypes on performance. As such, students pursuing more math-intensive domains, such as engineering or chemistry, may deem math ability as a highly valued attribute to their self-concept, and thus would be more likely to produce stronger stereotype threat effects than students pursuing psychology. Future research could consider recruiting students who are pursuing fields that require more mathematical abilities for success to see if threat effects are more salient to better understand whether the failed manipulation was a result of our demographics' experience with STEM or perceived lack of importance.

Provided that this study recruited college-aged psychology students (~20 years old), all of which had experience in a STEM field, future research could consider replicating this study using a younger demographic such as high school students to see what potential role age and experience may have on findings. It could be that younger individuals have less first-hand experience with STEM fields, and more so have not had as much experience with navigating STEM fields as a woman, and thus, may be more susceptible to the negative effects of stereotypes on performance. Further, it could be that younger demographics' academic identities are more malleable (lower self-esteem) than older demographics, and thus, are more likely to experience performance decrements when threat is activated. In fact, research has found that self-esteem tends to increase once an individual reaches young adulthood (i.e., Erol & Orth, 2011; Wagner, Ludtke,

Jonkemann, & Trautwein, 2013), which is thought to begin around the age of 18. This boost in self-esteem could serve as an additional protective barrier against stereotype threat effects, making older individuals less susceptible to these effects as shown in the current study. Future research could consider recruiting a younger demographic and/or consider including a self-esteem questionnaire to gauge what effect these factors may have on stereotype threat effects to better understand when intervention is most crucial.

Another limitation that may have led to our threat manipulation failing could be the method in which data was collected – participants completed the study in a single-session, online, using their own devices (outside of a lab), and thus, the study lacked rigorous experimental control. Provided that this study was not conducted in the presence of a researcher, in a controlled laboratory environment, confounding variables within the participants' environment could have hindered the ability to find conditional differences. It could be that findings would be more pronounced if conducted in a more controlled, laboratory setting, in order to account for any interference. Future directions could consider other means of data collection, such as in-person, to better understand the role that our means of data collection may have had on our results. Additionally, it could be that the format of the stereotype threat manipulation was not as effective as it could have been. Participants were simply told that research has found that men are superior at math abilities relative to women and that they would solve math problems. While threat was indeed activated, perhaps motivating performance, there was nothing really on the line for participants if their performance was poor. One way to further motivate performance could be by telling participants that their scores on the math task would be available for others to see and compare. This could further push participants to success, especially

those within the threat condition, in fear that their poor performance would confirm the negative stereotype about their gender group.

In the current study, the threat manipulation was presented within the instructions of the task in which participants did not have to actively engage with the manipulation (outside of answering a manipulation check). Important to note is that 166 participants were excluded for failing the manipulation check, and thus the salience of stereotype threat in the current study may have been too weak to promote performance deficits. Perhaps a threat manipulation that required participants' active engagement would be a more effective way of activating stereotype threat. One alternative could be asking participants to write about the threat manipulation to some extent (i.e., "what experiences have you had with gender differences in math abilities?"). Forcing participants to relate the manipulation back to their own lives, in a meaningful way, could lead to more pronounced findings including the replication of the basic stereotype threat effect.

Further, important to note is that participants exhibited moderate performance on the math problem-solving, regardless of condition. The average score was 5.23 (out of 10; $SD = 2.70$) and spanned the full range from 0 to 10, suggesting that the posterior probability problems were especially difficult for some participants to solve. While the intention was to challenge participants, it could be that these problems were too cognitively demanding of participants and thus, could have led to more participants disengaging from the task, which could have served to attenuate condition-level performance differences. Additionally, students tend to have more difficulty grasping probability theory compared to other mathematical skills (such as arithmetic or algebra)

and thus future research could select problems from domains of math that are more universally understood than probability theory (e.g., Weber, Binder, & Krauss, 2018).

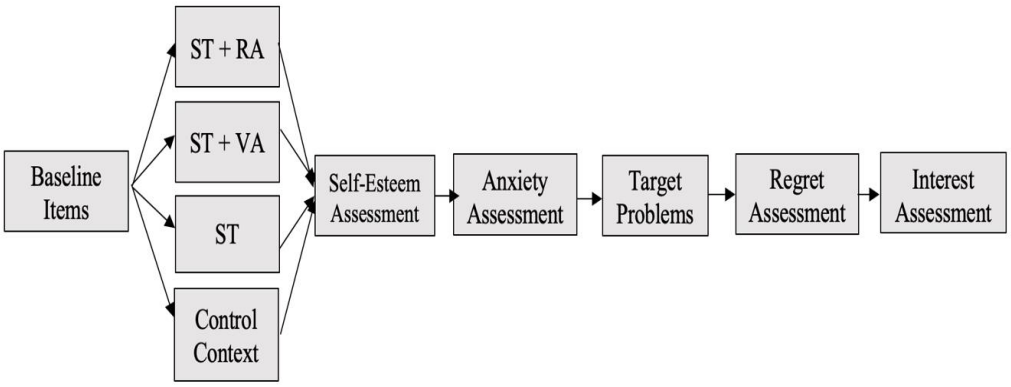
While this study looked at the effects of self-affirmation on math performance, future research should look to see if affirmation is universally beneficial for non-math abilities. For example, follow-up research could look at the effects of self-affirmation on other academic abilities aside from math, such as reading comprehension. It could be that affirming one's abilities, in turn, could act as a performance boost in all intellectual/academic abilities. Further, while this study looked at the effects of self-affirmation on boosting academic abilities, math performance, future research should look to see if the benefits of self-affirmation can be applied more broadly, in non-academic abilities (i.e., physical abilities). It could be that providing affirmation provides a confidence boost that could result in performance boosts, regardless of the domain. This future direction is especially worthy of investigation provided that affirmation is a relatively low-cost remedy to boost performance.

5.2 Conclusion

The current study contributes to existing research on stereotype threat and self-affirmation by testing the differential benefits of two types of affirmation on reducing stereotype threat effects on women's math performance. The most novel, and pronounced finding was that relational affirmation was most beneficial for boosting women's performance. While this study was unable to replicate stereotype threat effects, future research should further investigate the universal benefits of self-affirmation.

APPENDIX A – FIGURES

Figure 1. *Visual Schematic of the Design and Procedure*.



APPENDIX B – TABLES

Table 1. *Table of Means, SDs, and Correlations.*

	Threat Relational Affirmation	Threat Verbal Affirmation	Threat No Affirmation	No Threat No Affirmation
Sample Size	68	81	81	78
Target Performance	5.79 (2.83)	5.23 (2.62)	5.15 (2.86)	4.83 (2.45)
Winsorized Target Performance	5.68 (2.66)	5.16 (2.49)	5.07 (2.66)	4.78 (2.34)
State Self-Esteem	3.32 (.69)	3.29 (.76)	3.19 (.68)	3.37 (.65)
Anxiety	3.57 (1.79)	3.72 (1.83)	3.94 (1.71)	3.83 (1.57)
Regret	2.77 (.72)	2.82 (.69)	2.78 (.76)	2.82 (.61)
STEM Interest	11.18 (5.05)	11.33 (5.36)	11.81 (5.05)	12.51 (6.09)
Non-STEM Interest	12.37 (6.45)	13.12 (5.94)	13.15 (5.91)	14.21 (5.60)

Table 2. *Table of Correlations.*

	1	2	3	4	5
1. Self-Esteem	-	-.42**	-.26**	.04	-.10*
2. Anxiety	-.42**	-	.20**	-.03	.11*
3. Regret	-.26**	.20**	-	-.05	.08
4. STEM Interest	.04	-.03	-.05	-	.51**
5. Non-STEM Interest	-.10*	.11*	.08	.51**	-

*$p < .05$.
**$p < .01$.

APPENDIX C FOOTNOTES

1 I analyzed state self-esteem by subscale as well, given that performance and social self-esteem could plausibly have been higher in the relational affirmation condition due to higher performance on the math test and reflecting on having better relational abilities. However, there was no effect of condition on levels of performance self-esteem, $F(3,304) = .970$, $p = .445$, $\eta^2_p = 0.004$. Further, there was no effect of condition on social self-esteem, $F(3,304) = 1.65$, $p = .179$, $\eta^2_p = 0.016$, or appearance self-esteem, $F(3,304) = .187$, $p = .905$, $\eta^2_p = 0.002$. Thus, levels of self-esteem (regardless of type) did not vary by condition.

2 To account for the large variance in responses on the 10-item math task (M = 5.23, SD = 2.70), I ran an outlier analyses in which target performance scores were Winsorized where scores below 1 were changed to a 1 (n = 5) and scores above 9 were changed to a 9 (n = 26). See Table 1 for descriptive statistics. We then reconducted the above analyses.

First, I conducted a one-way ANCOVA, controlling for differences in baseline math performance, with math performance as the dependent measure and condition (4 levels) as the independent variable. This analysis was again non-significant, $F(3, 303) = 1.59$, $p = .193$, $\eta^2_p = .015$. I then conducted two contrasts to determine if specific mean level differences emerged, independent of the non-significant omnibus analysis. In the first contrast, I simultaneously compared math performance in the two affirmation conditions to the two non-affirmation conditions. Again, a marginally significant effect emerged, $t(304) = 1.92$, $p = .066$, $d = .79$. In the second contrast, I conducted a more focused contrast in which I compared the relational affirmation condition simultaneously

to the other three conditions. A marginally significant effect emerged, $t(304) = 1.92$, $p = .056$, $d = .39$. Thus, reported findings are not due to extreme scores on the math test.

APPENDIX D MATERIALS

Example Probability Problem Including a Prevalence Problem (Question 1) and Predictive Value Problem (Question 2)

Problem Scenario: Breathalyzers

Imagine you work in a police department. Your department often uses Breathalyzers to test whether drivers are driving under the influence of alcohol. Based on previous cases in which a person's sobriety was later verified, you know the following:

	Positive Breathalyzer Test (Indicates drunkenness)	Negative Breathalyzer Test (Does not indicate drunkenness)
Sober Driver	A **150**	B **750**
Drunk Driver	C **75**	D **25**

1. **Based on this table, how likely is it that a driver is drunk?**
   a. 0.15
   b. 0.50
   c. 0.10
   d. 0.75

2. **Based on this table, how likely is it that a driver with a positive Breathalyzer test is actually drunk?**
   a. 0.66
   b. 0.95
   c. 0.30
   d. 0.33

State Self-esteem Assessment

You will now complete a personality questionnaire designed to measure what you are thinking at this moment. There is, of course, no right answer for any statement. The best answer is what you feel is true of yourself at the moment. Be sure to answer all of the items, even if you are not certain of the best answer. Again, answer these questions as they are true for you **RIGHT NOW** before beginning math problem-solving.

Not at all	A little bit	Somewhat	Very much	Extremely
□	□	□	□	□

1. I feel confident about my abilities.
2. I am worried about whether I am regarded as a success or failure. (r)
3. I feel frustrated or rattled about my performance. (r)
4. I feel that I am having trouble understanding things that I read. (r)
5. I feel that others respect and admire me.
6. I feel self-conscious. (r)
7. I feel as smart as others. (r)
8. I feel displeased with myself. (r)
9. I feel good about myself.
10. I am worried about what other people think of me. (r)
11. I feel confident that I understand things.
12. I feel inferior to others at this moment. (r)
13. I feel concerned about the impression I am making. (r)
14. I feel that I have less scholastic ability right now than others. (r)
15. I feel like I'm not doing well. (r)
16. I am worried about looking foolish. (r)

Regret Assessment

Please read the statements below and select an option for each line.

Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
□	□	□	□	□

1. I regret my performance on the problem-solving.
2. My identity feels threatened based on my performance on the problems.
3. I feel confident after completing the problems.
4. I could have done better on the problem-solving.
5. I did as good as I could have possibly done on the problem-solving.
6. I feel insecure after completing the problems.
7. I am satisfied with my performance.
8. My performance is a result of my gender.
9. My performance has nothing to do with my gender.
10. My performance is a result of my effort.
11. My performance has nothing to do with my effort.
12. I am disappointed with my performance.
13. I am proud of my performance.
14. I am ashamed of my performance.
15. I would feel uncomfortable sharing my score with members of my own gender.
16. I would feel uncomfortable sharing my score with members of the opposite gender.
17. I would feel comfortable sharing my score with members of my own gender.
18. I would feel comfortable sharing my score with members of the opposite gender.

Interest Assessment

Please read the questions below and select an option for each line.

1 – Not at all, 4 – Neutral, 7 – Very much

1. How in support are you of being required to take **an additional math course (e.g., statistics, algebra)**?
2. How in support are you of being required to take **an additional science course (e.g., biology, anatomy)**?
3. How in support are you of being required to take **an additional technology course (e.g., computer science, web development)**?
4. How in support are you of being required to take **an additional engineering course (e.g., physics, geometry)**?
5. How in support are you of being required to take **an additional history course (e.g., world history, U.S. history)**?
6. How in support are you of being required to take **an additional english course (e.g., creative writing, world literature)**?
7. How in support are you of being required to take **an additional art course (e.g., painting, sculpture)**?
8. How in support are you of being required to take **an additional physical education course (e.g., yoga, cycling)**?

Consent Form

University of Southern Mississippi Informed Consent Statement For
Differences in Problem Solving

You are invited to take part in a research study conducted by Alicia Macchione in the School of Psychology. Any questions or concerns regarding this research may be directed to Alicia Macchione (Alicia.Macchione@usm.edu). This project and this consent form have been reviewed by the Institutional Review Board, which ensures that research projects involving human participants follow federal regulations. Any questions or concerns about your rights as a research participant should be directed to the Chair of the Institutional Review Board, The University of Southern Mississippi, Box 5125, Hattiesburg, MS 39406, (601) 266-5997.

Research Overview.
This study is interested in how different individuals solve math problems. In this study, we will ask you to solve a series of math problems and respond to some writing prompts. Then, we will have you respond to a personality inventory and some demographics questions. Based on pre-testing, this study should take you no more than 30 minutes to complete if you complete this study undistracted.

Voluntary Participation.
You are free to discontinue your participation in this study at any time without penalty or loss of benefits. You may also freely decline to answer any of the questions asked of you.

Confidentiality.
The responses that you provide today will be kept completely confidential. At no time will your name or any other identifying information be associated with any of the data that you generate today. It will never be possible to identify you personally in any report of this research. Within these restrictions, results of the study will be made available to you upon request.

Risks & Benefits.
The risks associated with participation in this study are not greater than those ordinarily encountered in daily life, although you may feel mild emotional discomfort in various stages of the experiment. If you feel that you are distressed at any time while participating in this research, you should notify the researcher immediately. Your participation in this study does not guarantee any beneficial results. However, it will aid in your understanding of how psychological research is conducted as well as contribute to the general knowledge in the field.

Consent to Participate in Research.
In consideration of all of the above, I give my consent to participate in this research study. By checking below, you acknowledge that you have read and understood the above information, that you are 18 years of age or older and give your consent to

participate in our internet-based study. If you would like a copy of this consent form, please print it out now or email the experimenter for a copy.

If you consent to these procedures, please click the button labeled "**I agree to take part in this study.**" below and click the next arrow to begin. If you do not consent, please close the window now.

Debriefing Form

Thank you for participating in this study! We hope you found your experience interesting and enjoyable.

The purpose of this study is to investigate the effects of self-affirmation on mathematics problem-solving for women in different situations. There is some evidence suggesting that women *sometimes* perform worse on math tests than men. However, these differences are often small (and sometimes non-existent). Importantly, differences are more likely to happen when women are aware of the negative stereotype that "women are bad at math." This situation, in which women are aware of the negative stereotype, is called stereotype threat. This stereotype threat can be activated by telling women explicitly about the stereotype or by having them focus on their gender by marking "female."

In this study, some students solved problems in a "threat" condition in which we told them that women consistently underperform men on the target math task. Other students solved problems in a "non-threat" condition in which we told them that women and men consistently perform similarly on the target math task. We wanted to assess whether self-affirmation effects learning differently for women in threat vs. non-threat situations. Affirming one's values or skills, such as relational or verbal ability, can have a positive effect on learning and performance. We predicted that affirmation would help mitigate against the negative effects of stereotype threat on math performance.

For more information on stereotype threat, you can check out this website: http://www.reducingstereotypethreat.org/. It was created by social psychologists as a resource for people who are interested to learn more about stereotype threat and ways to combat it.

Data collected here will give us insight into how to improve mathematics learning and how to stop stereotypes from affecting math performance. If for any reason you do not want your responses to be included in the set of data that will be analyzed for this study, please contact an experimenter immediately: Alicia Macchione (Alicia.Macchione@usm.edu). Please also reach out if you have any questions or concerns. Thank you again for your participation.

APPENDIX E IRB APPROVAL LETTER

Office *of*
Research Integrity

THE UNIVERSITY OF SOUTHERN MISSISSIPPI

118 COLLEGE DRIVE #5125 • HATTIESBURG, MS | 601.266.6576 | USM.EDU/ORI

NOTICE OF INSTITUTIONAL REVIEW BOARD ACTION

The project below has been reviewed by The University of Southern Mississippi Institutional Review Board in accordance with Federal Drug Administration regulations (21 CFR 26, 111), Department of Health and Human Services regulations (45 CFR Part 46), and University Policy to ensure:

- The risks to subjects are minimized and reasonable in relation to the anticipated benefits.
- The selection of subjects is equitable.
- Informed consent is adequate and appropriately documented.
- Where appropriate, the research plan makes adequate provisions for monitoring the data collected to ensure the safety of the subjects.
- Where appropriate, there are adequate provisions to protect the privacy of subjects and to maintain the confidentiality of all data.
- Appropriate additional safeguards have been included to protect vulnerable subjects.
- Any unanticipated, serious, or continuing problems encountered involving risks to subjects must be reported immediately. Problems should be reported to ORI via the Incident template on Cayuse IRB.
- The period of approval is twelve months. An application for renewal must be submitted for projects exceeding twelve months.
- FACE-TO-FACE DATA COLLECTION WILL NOT COMMENCE UNTIL USM'S IRB MODIFIES THE DIRECTIVE TO HALT NON-ESSENTIAL (NO DIRECT BENEFIT TO PARTICIPANTS) RESEARCH.

PROTOCOL NUMBER: IRB-20-302
PROJECT TITLE: Thesis: Self-Affirmation and Stereotype Threat in Math
SCHOOL/PROGRAM: Psychology
RESEARCHER(S): Alicia Macchione, Donald Sacco

IRB COMMITTEE ACTION: Approved
CATEGORY: Expedited

PERIOD OF APPROVAL: July 15, 2020

Michael Madson, Ph.D.
Institutional Review Board Associate Chairperson

REFERENCES

Andreano, J. M., & Cahill, L. (2009). Sex influences on the neurobiology of learning and memory. *Learning & Memory, 16*(4), 248–266. doi:10.1101/lm.918309

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice Hall.

Benbow, C. P., Lubinski, D., Shea, D. L., & Eftekhari-Sanjani, H. (2000). Sex differences in mathematical reasoning ability at age 13: Their status 20 years later. *Psychological Science, 11*(6), 474–480. doi:10.1111/1467-9280.00291

Brown, R. P., & Josephs, R. A. (1999). A burden of proof: Stereotype relevance and gender differences in math performance. *Journal of Personality and Social Psychology, 76*, 246-257. doi:10.1037/0022-3514.76.2.246

Brunstein, J., & Gollwitzer, P. (1996). Effects of failure on subsequent performance: The importance of self-defining goals. *Journal of Personality and Social Psychology. 70*. 395-407. doi:10.1037//0022-3514.70.2.395

Ceci, S. J., Williams, W. M., & Barnett, S. M. (2009). Women's underrepresentation in science: Sociocultural and biological considerations. *Psychological Bulletin, 135*, 218-61. doi:10.1037/a0014412

Cardwell, M. (1996). *Dictionary of Psychology*. Chicago IL: Fitzroy Dearborn.

Chalabaev, A., Stone, J., Sarrazin, P., & Croizet, J. C. (2008). Investigating physiological and self-reported mediators of stereotype lift effects on a motor task. *Basic and Applied Social Psychology, 30*(1), 18–26. doi:10.1080/01973530701665256

Cheryan, S., Master, A., & Meltzoff, A. N. (2015). Cultural stereotypes as gatekeepers: Increasing girls' interest in computer science and engineering by diversifying stereotypes. *Frontiers in Psychology, 6,* 49. doi:10.3389/fpsyg.2015.00049

Cohen, G. L., Aronson, J., & Steele, C. M. (2000). When beliefs yield to evidence: Reducing biased evaluation by affirming the self. *Personality and Social Psychology Bulletin, 26,* 1151–1164. doi: 10.1177/01461672002611011

Creswell, J. D., Dutcher, J. M., Klein, W. M. P., Harris, P. R., & Levine, J. M. (2013). Self-affirmation improves problem-solving under stress. *PLoS ONE 8*(5): e62593. doi:10.1371/journal.pone.0062593

Dasgupta, N., & Stout, J. G. (2014). Girls and women in science, technology, engineering, and mathematics: STEMing the tide and broadening participation in stem careers. *Policy Insights from the Behavioral and Brain Sciences, 1*(1), 21–29. doi:10.1177/2372732214549471

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. doi:10.1016/j.lindif.2015.12.018

Faul, F., Erdfelder, E., Lang, A., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods* 39, 175–191. doi:10.3758/BF03193146

Fein, S., & Spencer, S. J. (1997). Prejudice as self-image maintenance: Affirming the self through derogating others. *Journal of Personality and Social Psychology, 73,* 31–44. doi:10.1037/0022-3514.73.1.31

Fyfe, E. R., & Brown, S. A. (2018b). Task expectations influence learning from feedback. In T. T. Rogers, M. Rau, X. Zhu, & C. W. Kalish (Eds.), *Proceedings of the 40th Annual Conference of the Cognitive Science Society*, (pp. 396-401). Austin, TX: Cognitive Science Society.

Galinsky, A. D., Gruenfeld, D. H., & Magee, J. C. (2003). From power to action. *Journal of Personality and Social Psychology, 85*, 453-466.

Guinote, A. (2007a). Power affects basic cognition: Increased attentional inhibition and flexibility. *Journal of Experimental Social Psychology, 43*, 685-697.

Guinote, A. (2007b). Power and goal pursuit. *Personality and Social Psychology Bulletin, 33*, 1076-1087.

Harris, P. S., Harris, P. R., & Miles, E. (2017). Self-affirmation improves performance on tasks related to executive functioning. *Journal of Experimental Social Psychology, 70,* 281–285. https://doi.org/10.1016/j.jesp.2016.11.011

Heatherton, T. F. & Polivy, J. (1991). Development and validation of a scale for measuring state self-esteem. *Journal of Personality and Social Psychology, 60,* 895-910.

Hoffrage, U., Kurzenhauser, S., & Gigerenzer, G. (2005). Understanding the results of medical tests: Why the representation of statistical information matters. In R. Bibace, J. D. Laird, K. L. Noller, & J. Valsiner (Eds.), *Science and Medicine in Dialogue*. Westport: Praeger Publishers.

Hyde, J. S., & Linn, M. C. (1988) Gender differences in verbal-ability: a meta-analysis. *Psychological Bulletin, 104*(1), 53–69.

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,* 365–371. doi:10.1111/1467-9280.00272

Johnson, H., Barnard-Brak, L., Saxon, T., & Shen, M. (2012). An experimental study of the effects of stereotype threat and stereotype lift on men and women's performance in mathematics. *Journal of Experimental Education, 80,* 137-149. doi:10.1080/00220973.2011.567312

Kang, S. K., Galinsky, A. D., Kray, L. J., & Shirako, A. (2015). Power affects performance when the pressure is on: Evidence for low-power threat and high-power lift. *Personality & Social Psychology Bulletion, 41*(5), 726-35. doi: 10.1177/0146167215577365

Kimura, D. (1992) Sex differences in the brain. *Scientific American, 267*(3), 118–125.

Koole, S. L., Smeets, K., & Van Knippenberg, A., & Dijksterhuis, A. (1999). The cessation of rumination through self-affirmation. *Journal of Personality and Social Psychology. 77*. 111-125. doi:10.1037/0022-3514.77.1.111.

Korn Ferry Group [KFY]. (2016). Women outperform men in 11 of 12 key emotional intelligence competencies. Retrieved from https://www.kornferry.com/about-us/press/new-research-shows-women-are-better-at-using-soft-skills-crucial-for-effective-leadership

Levy, B. (1996). Improving memory in old age through implicit self-stereotyping. *Journal of Personality & Social Psychology, 71*(6), 1092-1107. doi:10.1037/0022-3514.71.6.1092

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

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

Master, A., & Meltzoff, A. (2017). Building bridges between psychological science and education: Cultural stereotypes, STEM, and equity. *Prospects. 46,* 1-20. doi:10.1007/s11125-017-9391-z

Mikulincer, M. (1996). Mental rumination and learned helplessness: Cognitive shifts during helplessness training and their behavioral consequences. In I. G. Sarason, G. R. Pierce, & B. R. Sarason (Eds.), *Cognitive interference: Theories, methods, and findings* (pp. 191-209). Mahwah, NJ: Erlbaum.

Nosek, B. A., Smyth, F. L., Sriram, N., Lindner, N. M., Devos, T., Ayala, A., et al. (2009). National differences in gender–science stereotypes predict national sex differences in science and math achievement. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(26), 10593–10597. doi:10.1073/pnas.0809921106

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

Pennington, C., Litchfield, D., McLatchie, N., & Heim, Derek. (2018). Stereotype threat may not impact women's inhibitory control or mathematical performance:

Providing support for the null hypothesis. *European Journal of Social Psychology. 49*. doi:10.1002/ejsp.2540

Rothwell, J. T. (2013). *The Hidden STEM Economy.* Metropolitan Policy Program. Retrieved from https://www.brookings.edu/research/the-hidden-stem-economy/

Rydell, R. J., & Boucher, K. L. (2017). Stereotype threat and learning. In J. M. Olson (Ed.), Advances in experimental social psychology: Vol. 56. *Advances in experimental social psychology* (p. 81–129). Elsevier Academic Press.

Rydell, R. J., Rydell, M. T., & Boucher, K. L. (2010). The effect of negative performance stereotypes on learning. *Journal of Personality and Social Psychology, 99,* 883–896. doi:10.1037/a0021139

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

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

Sherman, D. K. (2013). Self-affirmation: Understanding the effects. *Social and Personality Psychology Compass, 7*(11), 834–845. doi:10.1111/spc3.12072

Sherman, D. K., & Cohen, G. L. (2006). The psychology of self-defense: Self-affirmation theory. In M. P. Zanna (Ed.), *Advances in Experimental Social Psychology: Vol. 38,* (p. 183–242). Elsevier Academic Press. doi:10.1016/S0065-2601(06)38004-5

Shih, M., Pittinsky, T. L., & Ambady, N. (1999). Stereotype susceptibility: Identity salience and shifts in quantitative performance. *Psychological Science, 10*(1), 80-83. doi:10.1111/1467-9280.00111

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

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

Steele, C. M. (1988). The psychology of self-affirmation: Sustaining the integrity of the self. In L. Berkowitz (Ed.) *Advances in Experimental Social Psychology* (Vol. 21, pp. 261-302). New York: Academic Press.

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

Steele, C. M., & Liu, T. J. (1983). Dissonance processes as self-affirmation. *Journal of Personality and Social Psychology, 45*, 5-19. doi: 10.1037/0022-3514.45.1.5

Stewart, Fran. (2018). STEM and the local Economy: Do regions reap the benefits of a STEM-educated workforce? *Employment Research, 25(*1): 1-4. doi:10.17848/1075-8445.25(1)-1

Sweller, J., Van Merrienboer, J., & Paas, F. (1998). Cognitive architecture and instructional design. *Educational Psychology Review, 10*(3), 251-296. doi: 10.1023/a:1022193728205

Weber, P., Binder, K., & Krauss, S. (2018). Why can only 24% solve Bayesian reasoning problems in natural frequencies: Frequency phobia in spite of probability blindness. *Frontiers in Psychology, 9,* 1833. doi:10.3389/fpsyg.2018.01833

Walton, G., & Cohen, G. (2003). Stereotype lift. *Journal of Experimental Social Psychology*, *39*, 456–467. doi:10.1016/S0022-1031(03)00019-2

Wang, M.-T., Eccles, J. S., & Kenny, S. (2013). Not Lack of Ability but More Choice: Individual and Gender Differences in Choice of Careers in Science, Technology, Engineering, and Mathematics. *Psychological Science*, *24*(5), 770–775.

Whitson, J. A., Liljenquist, K. A., Galinsky, A. D., Magee, J. C., Gruenfeld, D. H., & Cadena, B. (2013). The blind leading: Power reduces awareness of constraints. *Journal of Experimental Social Psychology, 49,* 579-582.

Xie, Y., Fang, M., & Shauman, K. (2015). STEM education. *Annual Review of Sociology, 41*, 331–357. doi:10.1146/annurev-soc-071312-145659