

STEREOTYPE THREAT AND BLACK WOMEN'S MATH PERFORMANCE: DO
RACIAL IDENTITY AND EXPERIMENTER RACE MATTER?

BY

ARIELLE A. BROWN

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Educational Psychology
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2019

Urbana, Illinois

Doctoral Committee:

Professor Helen Neville, Chair
Associate Professor Mark Aber
Associate Professor Carla Hunter
Dr. Claytie Davis III, University of California, Berkeley

ABSTRACT

In this study, I examined the effect of stereotype threat on the math task performance among Black undergraduate women ($N = 103$). Similar to Davis et al. (2006), I explored if racial identity beliefs moderated test performance in the face of stereotype threat. Although there is growing research on stereotype threat, few stereotype threat studies have (a) included racial identity as a moderator of test performance or (b) studied the math performance of Black women with stereotype threat (Wright-Adams, 2014). The purpose of this study was to extend Davis et al.'s (2006) and Wright-Adams' (2014) research by exploring the race of the experimenters (i.e., White or Black woman) on the relationship between stereotype threat and math performance. Among Black women college students, findings from analyses of covariance (ANCOVAs) revealed no significant effect of stereotype threat on math performance. Lastly, racial identity did not significantly moderate the association between stereotype threat and math task performance. However ACT math scores were positively and significantly linked to performance on the math task. Implications of findings are discussed.

Keywords: Stereotype threat, Black, African American, women, math, racial identity

ACKNOWLEDGEMENTS

Personally and firstly, I would like to thank God. Without His strength, love, and wisdom, this journey and completion would have not been possible for me. I would also like to thank my amazing, hard-working, and supportive parents—Aaron and Darlene Brown. I would not be where I am without you two. Thank you for all of the sacrifices you have made on my behalf. You have always wanted more for me than what I wanted for myself and I owe you more than you will ever realize. Moreover, I want to thank my precious and strong sister, Deishall Brown and my fantastic nephew, Skylar, who add to my joy daily and who make me feel more than special. My great friends—Erineka and Liz, your love and loyalty are immeasurable.

Personally *and* professionally, I would like to acknowledge and thank Dr. Helen Neville, my brilliant advisor and chair, who has never given up on me and who has played a pivotal role in all of my research efforts. Your private acts of love, such as funding this project, have never gone unnoticed and I am so appreciate of everything you do for me. I am truly blessed to be your advisee. A special thanks also goes to the wonderful women, Cassandra Colbert and Dawn Brown, who worked as experimenters for this study. Your sacrificed time and effort will never be forgotten. I also give my gratitude to the undergraduate women who worked as research assistants for my study. You all made this process much smoother than I could imagine. Finally, I would like to extend my gratitude to my committee members, Dr. Mark Aber, Dr. Claytie Davis III, and Dr. Carla Hunter. You all were very instrumental during this process, and I am fortunate to have a flexible and committed group like you all. Thank you again everyone!

TABLE OF CONTENTS

CHAPTER 1: INTRODUCTION.....	1
CHAPTER 2: LITERATURE REVIEW.....	5
CHAPTER 3: METHOD	27
CHAPTER 4: RESULTS	34
CHAPTER 5: DISCUSSION.....	43
REFERENCES	50
APPENDIX A: DEMOGRAPHIC SURVEY.....	59
APPENDIX B: AFRICAN AMERICAN CULTURAL ENGAGEMENT.....	62
APPENDIX C: AFRICAN AMERICAN WOMEN CULTURAL ENGAGEMENT.....	63
APPENDIX D: CONTROL GROUP (LOW THREAT)	64
APPENDIX E: EXPERIMENTAL SCRIPT (HIGH THREAT).....	65
APPENDIX F: DEBRIEFING SCRIPT.....	66
APPENDIX G: EXPERIMENTER PROTOCOL.....	68
APPENDIX H: EXPERIMENTER CHECKLIST.....	70
APPENDIX I: DESCRIPTIVE STATISTICS.....	71
APPENDIX J: CORRELATION MATRIX.....	72
APPENDIX K: MAIN ANALYSES WITH LISTWISE DELETION METHOD FOR MISSING ACT SCORES.....	73
APPENDIX L: ADDITIONAL MEASURES.....	75
APPENDIX M: IRB APPROVAL.....	81

CHAPTER 1: INTRODUCTION

Stereotype threat phenomenon occurs when individuals are in a situation where they feel threatened about confirming a stereotype (usually negative) about their social group (Steele & Aronson, 1995). Stereotype threat affects members of marginalized groups, or those with negative stereotypes associated with membership of a group, such as Black individuals and women (Steele, 2010; Steele & Aronson, 1995). To illustrate, African Americans have been stereotyped as less intelligent than their White counterparts, and studies have shown that this stereotype has affected their cognitive tasks performance. In early stereotype threat studies, Black students performed significantly worse on difficult GRE items compared to White students after they were reminded of their race prior to taking the test. When they were not reminded of their race, they performed better than their White peers (Steele & Aronson, 1995). A growing body of research indicates stereotype threat among African Americans is connected to poorer test and academic performance (Owens & Massey, 2011; Steele, 2010) and distressed cognitive processes after being exposed to a stereotype threat cue (Steele & Aronson, 1995). Other social identity groups can experience stereotype threat as well. For example, women's math performance decreased after they were given a stereotype threat cue (Nguyen & Ryan, 2008; Spencer, Steele, & Quinn, 1999; Walton & Cohen, 2003), and the effects were more severe for women with moderate or high levels of math identification (Nguyen & Ryan, 2008; Walton & Cohen, 2003).

There are a few studies on stereotype threat and math performance, mostly suggesting that women's performance is significantly hindered when they are reminded of their gender prior to completing math items (Nguyen & Ryan, 2008; Spencer et al.,

1999; Walton & Cohen, 2003). Although there are numerous studies about stereotype threat, there are gaps in the research in exploring race-related moderators. Only a few studies have (a) explicitly studied the math performance of African American women in the face of stereotype threat, (b) included racial identity attitudes as a moderator of test performance, and (c) no stereotype threat study has included racial identity attitudes and the race of the experimenter as moderators for African American women in one study.

In efforts to minimize the effects of stereotype threat, scholars have investigated if factors such as racial identity beliefs (i.e., how one evaluates their racial group) moderates the effects of stereotype threat on academic tasks such as math performance. Most racial identity studies in this area focus on the beliefs one has about their racial group membership; there is little consideration of one's behaviors reflecting these beliefs. In addition to racial identity attitudes, this study also explores if a *behavioral* component of racial identity, cultural engagement, is connected to stereotype threat performance. Cultural engagement refers to individuals' participation in activities connected to their racial group (Brannon, Markus, & Taylor, 2015). In this study, racial identity behavior is also defined as participating in activities related to one's racial group and is one component of racial identity. Higher levels of cultural engagement are related to better academic outcomes in general, such as academic fit and identification, higher GPA, and academic persistence (Brannon et al., 2015; Mendoza-Denton, Downey, Purdie, Davis, & Pietrzak, 2002). To date, however, researchers have not examined if cultural engagement buffers the negative influence of stereotype threat (Brannon et al., 2015). This notion of racial identity behavior coupled with racial identity beliefs could help provide a more

complex understanding of the association between stereotype threat and math performance.

Drawing on stereotype threat and racial identity theories, I used an experimental design to explore if race-related variables moderated the link between stereotype threat and math performance. This study takes an intersectionality approach, as it acknowledges the potential effect of stereotype threat on Black women who tend to experience various and intersecting forms of marginalization and oppression. Building on the growing literature, the following four research questions directed the study.

Does stereotype threat matter for Black women’s math performance?

- Research Question 1: Do participants in the low stereotype threat condition perform better on a math task than students in the high stereotype threat condition? On the basis of previous research, I hypothesized that African American women in the low stereotype threat condition will perform better on the math task than students in the high stereotype threat condition.

Does race of experimenter matter for Black women’s math performance?

- Research Question 2: Does the race of women test administrators moderate the effects of stereotype threat on a math task among African American women? This question is exploratory, as no study has specifically examined if experimenter race moderates the association between stereotype threat and math task performance for Black women.

Does racial identity matter?

- Research Question 3: Do racial identity *beliefs* moderate the effects of stereotype threat on a math task among African American women? Specifically, I tested if

participants in the low threat condition with higher Internalization Multiculturalist Inclusive (IMCI; Vandiver et al., 2001) attitudes performed better on the math task than participants in the high threat condition with lower IMICI attitudes. The IMCI subscale (Vandiver et al., 2001) was only used in the current study because Davis and colleagues (2006) found higher scores on the Internalization subscale (Racial Identity Attitudes Scale—revised; RIAS; Helms & Parham, 1990) to be associated with increased performance in the low threat condition. Moreover, this investigation takes a strengths-based approach in assessing if more positive and internalized racial identity beliefs act as buffers to stereotype threat effects.

- Research Question 4: Do racial identity *behaviors* (i.e., cultural engagement) moderate the effects of stereotype threat on a math task among African American women? Specifically, I tested if participants in the low threat condition with higher cultural engagement perform better on the math task than participants in the high threat condition with lower cultural engagement. This research question is also exploratory, as there are no current studies that examine the moderating role of racial identity behaviors in the stereotype threat – math performance link.

CHAPTER 2: LITERATURE REVIEW

In this chapter, I define stereotype threat phenomenon and provide a summary of its effect on the testing and academic performance of African American college students. Subsequently, I discuss the effects of stereotype threat on math performance among women, and then specifically African American women. Last, potential moderators of stereotype threat are examined, such as the race of the experimenter, and racial identity beliefs and behaviors.

STEREOTYPE THREAT

Social psychologists describe stereotype threat as being at risk of confirming a negative stereotype regarding one's social group (Steele & Aronson, 1995). It also has been broadly described as having to face certain group generalizations and circumstances, often threatening and discriminatory, solely because the individual has a particular social identity (Steele, 2010). The effects of stereotype threat can manifest in the following conditions: (1) an individual is familiar with a negative stereotype about their group; (2) the threat is triggered by the recognition of a negative stereotype in a performance situation (e.g., test, presentation, athletic performance); (3) the individual identifies with the domain (e.g., identifies with math); and (4) the individual fears confirming the stereotype and their performance is hindered. Steele (2010) argued stereotype threat yields a myriad of negative unforeseen circumstances, and in his earlier work, he focused on discovering how those situations affected African American students academically.

There are multiple ways of inducing stereotype threat and researchers use diverse methods to assess its effects among Black individuals. One of the ways researchers have invoked stereotype threat in experimental studies is through scripts. Such scripts include

diagnosticity, or a statement indicating that one's performance on the task reflects their abilities or characteristics. To illustrate, phrases such as, "verbal ability or verbal intelligence" (Davis et al., 2016, p. 407) have been used in diagnostic scripts. In other experimental studies, some scripts primed racial group membership, as participants were asked to disclose their race before the test or task was administered (Ho & Sidanius, 2010; McFarland et al., 2003).

Some experiments have invoked stereotype threat with other methods such as surveys that measure stereotype biases, stereotypical ideologies, and stereotype vulnerability (Cromley et al., 2013; Owens & Massey, 2011). Vulnerability to race stereotype bias has been measured using items such as, "I believe that if I perform well on a test, the professor will attribute my good performance to my race" (Cromley et al., 2013, p. 251), and stereotypical ideologies have been measured with items such like, "On a scale of 0 (lazy) to 6 (hardworking), do members of your own racial group tend to be lazy or hardworking?" (Owens & Massey, 2011, p. 153).

In sum, stereotype threat experimental studies have included laboratory methods (Davis et al., 2006; Ho & Sidanius, 2010; Steele, 2010; Steele & Aronson, 1995; Taylor & Walton, 2011) and non-laboratory approaches (Cromley et al., 2013; Owens & Massey, 2011). Stereotype threat intervention studies have also been conducted (Cohen, Garcia, Apfel, & Master, 2006; Simmons, 2013; Steele, 1997; Steele et al., 1998; Walton & Cohen, 2007). The diversity in foci and methods in stereotype research can serve as a limitation in making conclusions and replicating studies especially if researchers do not consider similarities and differences among studies.

STEREOTYPE THREAT AND AFRICAN AMERICANS

Stereotype threat among African Americans has been connected to poorer test performance and grade point average (Craemer & Orey, 2017; Davis et al., 2006; Ho & Sidanius, 2010; Steele & Aronson, 1995). Generally, findings suggest after Black participants are exposed to a stereotype threat cue, they perform worse on cognitive tasks compared to when they are not exposed to such a threat. Steele and Aronson (1995) were the first researchers to explicitly test stereotype threat. They examined whether African American participants performed worse than White participants on a challenging verbal cognitive task when the test administrator stated the test measured intellectual verbal ability (diagnostic condition) before administering the test. However, when the participants were told that the test was being used to understand the "psychological factors involved in solving verbal problems..." (p. 799), the African American participants performed equally as well as the White participants. These findings suggest an *achievement ceiling* or limitation in academic achievement that could possibly exist for Black students in the face of threat.

Similar to Steele and Aronson's (1995) results, many scholars found that when manipulating testing conditions by creating diagnostic (e.g., an indicator of intelligence) or non-diagnostic instructions (e.g., an indicator of problem solving) in laboratory settings, stereotype threat negatively and significantly influenced Black students' academic performance (e.g., Davis et al., 2006; Ho & Sidanius, 2010; Nguyen & Ryan, 2008; Steele, 2010; Taylor & Walton, 2011). These debilitating effects of stereotype threat have been attributed to memory recall difficulties due to threat exposure. For example, Taylor and Walton (2011) found that participants' ability to recall words from a

word list in a threatening environment was significantly worse than those who were not exposed to a threat. Notably, when Black participants learned a word list in a non-threatening environment and recalled the words in a non-threatening environment, they actually performed *better* than the White participants in the study.

Three meta-analyses found negative effects of stereotype threat on a range of outcomes (Nguyen & Ryan, 2008; Walton & Cohen, 2003; Walton & Spencer, 2009). While these effect sizes for the most part were small (around $d = .26$), Nguyen and Ryan (2008) illustrated the implications of stereotype threat: “If a minority student took the SAT and his or her true cognitive ability were at the national mean level, he or she might underperform by about 50 points due to subtle stereotype threat cues” (p. 1330).

Cognitive processes indicating distress have also been associated with stereotype threat among Black students (Steele, 2010; Steele & Aronson, 1995). For example, African Americans in diagnostic experimental conditions indicated more stereotypic thoughts, greater concerns about their intellectual ability, avoided racially stereotypical preferences more, selected more excuses for their performance, and they were more apprehensive about indicating their race on their test (Steele & Aronson, 1995). These findings highlight how the consideration and avoidance of Black racial stereotypes are involved with test-taking under threatening conditions. Students in these studies not only cognitively processed negative stereotypes, they attempted to avoid them through other cognitive strategies and not stating their racial group membership.

Although most studies indicate impaired performance, stereotype threat has also been associated with null effects (Cromley et al., 2013; McFarland, Lev-Arey, & Ziegert, 2003; Oliver et al., 2017; Steele, 2010). For example, Cromley and colleagues (2013)

found Black participants' scores on stereotype bias items did not change over time and they were not related to chemistry and biology course grades and retention in STEM. However, Black students had higher stereotype bias than White students in the beginning of the school year. It should be noted Cromley et al. measured one's *vulnerability* to threat; the researchers did not actually expose the students to threat as in other stereotype threat studies. It should also be noted that grades and retention were assessed over time, whereas other studies test performance shortly after threat exposure.

Findings from a few studies suggest invoking stereotype threat leads to an *increase* in academic and test performance (Owens & Massey, 2011; Wright-Adams, 2013). Owens and Massey (2011) found that Black college students with internalized stereotype threat, meaning they believed the negative stereotypes about their group, had better academic performance. The authors argued students wanted to prove they were not like others who were negatively viewed in their racial group. Similarly, in Wright-Adams (2014) study, the students in the high threat condition (i.e., diagnostic in which race and gender were mentioned) performed better than the students in the low threat condition (i.e., non-diagnostic in which race and gender were not mentioned) in a group testing setting. She posited that these findings were due to situational factors, such as including an African American experimenter and mostly African American women participants.

There are also mixed results from intervention studies. Although most intervention studies have yielded positive effects (Cohen, Garcia, Apfel, & Master, 2006; Steele, 1997; Steele et al., 1998; Walton & Cohen, 2007), a few studies have not (Simmons, 2013). Intervention research in this area often seeks solutions to reduce stereotype threat in academic settings by using self-affirmation activities to promote self-

integrity. African American students in intervention groups using self-affirmation strategies performed significantly better academically than students in non-intervention groups (Cohen et al.'s, 2006; Steele, 1997; Steele et al., 1998; Walton & Cohen, 2007).

There are a few studies in which participation in self-affirmation interventions did not show promising results. Simmons (2013) found null results among African American high school students who were at risk for academic failure. Students in the self-affirmation group were told to select one to three values that were *most* important to them and they explained why, whereas the other-affirmation group was told to select one to three values that were *least* important to them and they explained how it might be important to others. The results showed that students in the self-affirmation group did not have significantly higher winter grades or school engagement than the students in the other-affirmation group. These inconsistent findings with previous intervention literature could be attributed to the sample; the participants were at risk for academic failure and the intervention may have not been sufficient in changing academic behaviors.

In sum, there is mounting research suggesting stereotype threat among Black Americans is connected to poorer performance, particularly, immediately after exposure to a group-based stereotype (Davis et al., 2006; Ho & Sidanius, 2010; Nguyen & Ryan, 2008; Owens & Massey, 2011; Steele, 2010; Steele & Aronson, 1995; Taylor & Walton, 2011; Walton & Cohen, 2003; Walton & Spencer, 2009). However, a few studies suggest there is no relation between stereotype threat and academic performance (Cromley et al., 2013; McFarland, Lev-Arey, & Ziegert, 2003; Steele, 2010). Such studies tend to assess longer-term performance such as grades. Self-affirming methods in intervention research has also yielded mixed findings. A number of studies show that academic performance

could be increased with appropriate intervention methods for some, particularly those who are not at-risk for academic failure (Cohen et al., 2006; Simmons, 2013; Steele, 1997; Steele et al., 1998; Walton & Cohen, 2007). The type of intervention, the number of ethnic minorities in the setting, and what is used as the dependent variable (e.g., GPA vs. standardized test score) can account for the difference in findings. While effect sizes in meta-analyses and intervention studies are relatively small, there is evidence indicating threatening conditions decrease Black individuals' academic performance. The threatening conditions can be attributed to the widely known anti-Black intellectualism stereotype, and this debilitating stereotype is relevant to other marginalized groups such as Latinx individuals and women in STEM fields, regardless of racial or ethnic background (Owens & Massey, 2011; Steele, 2010). Specifically, women in STEM fields can experience a lack of belonging due to women's math inferiority stereotypes, which could affect their performance (Good, Rattan, & Dweck, 2012; Murphy, Steele, & Gross, 2007).

WOMEN AND STEREOTYPE THREAT

Stereotype threat among women has been associated with poorer math performance and math identification (Nguyen & Ryan, 2008; Picho, Rodriguez, & Finnie, 2013; Spencer, Steele, & Quinn, 1999; Walton & Cohen, 2003). Women who least identified with math were less negatively affected by stereotype threat compared to those who were highly or moderately identified with math (Nguyen & Ryan, 2008). Moreover, research shows that White women and racial/ethnic minority women and men all yield different performances from one another in stereotype threat situations (Nguyen & Ryan, 2008). Lastly, there is a lack of clarity regarding participants' racial and gender

identification in the stereotype threat literature.

Stereotype threat for women is related to inferior math performance compared to their male counterparts across studies, and the gap is largest when the math task is difficult (Nguyen & Ryan, 2008; Spencer et al., 1999; Walton & Cohen, 2003). Spencer and colleagues (1999) found that when women participants were reminded of their gender (e.g., “the test had shown gender differences in the past”), their scores on the math task were significantly lower than the control group. Oddly, the race of the participants was not provided. Women overall performed significantly worse than men on difficult math GRE questions in threatening conditions, but there was no difference in performance when the test was easy. It has been estimated that women’s performance is usually hindered about 19 to 21 points on the SAT math test due to stereotype threat, which highlights stereotype threat effects for women on standardized test performance as well as classroom math tasks (Walton & Spencer, 2009).

Moderate and high math identification has been associated with poorer math performance among women (Nguyen & Ryan, 2008; Walton & Cohen, 2003). Nguyen and Ryan (2008) found that White women and racial and ethnic minority college and high school students’ performance was hindered more when they moderately identified with math. That is, women who held moderate levels of math identity were more susceptible to stereotype threat. These findings are inconsistent with previous stereotype cross-sectional studies and meta-analyses, which have found that highly math-identified women were more susceptible to the effects of stereotype threat than those with lower math identification (e.g., Steele, 2010; Walton & Cohen, 2003). Nguyen and Ryan (2008) attributed this unexpected domain identification finding to the inconsistent and unclear

operationalization in stereotype threat studies. They also considered that when assessing domain identity, many studies did not directly examine the concept and mostly included highly domain-identified participants (e.g., math majors), which truncates the spectrum and understanding of domain identification.

Stereotype threat research with White women and ethnic minority men and women has yielded significant differences in performance and stereotype threat cues (Nguyen & Ryan, 2008). Findings across multiple stereotype threat studies have shown that the stereotype threat effects were more debilitating for “ethnic minorities” than for “women” test takers, and this effect was exacerbated when the test was highly difficult (Nguyen & Ryan, 2008). The researchers posited that the differences could be explained by inconsistencies in the operationalization of “test difficulty” in stereotype threat studies (Nguyen & Ryan, 2008). Highlighting an additional difference, subtle cues were more harmful for “women” while moderate stereotype cues affected “ethnic minority” test takers the most (Nguyen & Ryan, 2008). The exact number of women of color in these stereotype threat meta-analyses and individual studies have not been reported (e.g., Nguyen & Ryan, 2008; Spencer, Steele, & Quinn, 1999), limiting the understanding of how stereotype threat was shown to affect ethnic minority women across studies. Research that explicitly addresses the effects of stereotype threat with a racial-gendered lens is needed.

In sum, stereotype threat among women has been connected to poorer math performance (Nguyen & Ryan, 2008; Spencer et al., 1999; Walton & Cohen, 2003), moderate and high levels of math identification (Nguyen & Ryan, 2008; Walton & Cohen, 2003), and better performance and more sensitivity to subtle stereotype threat

cues compared to ethnic minorities (Nguyen & Ryan, 2008). Although the influence of stereotype threat on women's math performance has been studied extensively, most of the women participants in published studies have been White or the race/ethnicity of the participants have not been provided. There is a gap in the literature in understanding the effects of stereotype threat on math performance for women of color, particularly Black women.

AFRICAN AMERICAN WOMEN AND STEREOTYPE THREAT

The stereotype threat literature regarding women and math performance has often neglected the influence of the intersection of gender and race. The earlier literature of stereotype threat revealed that African Americans' (regardless of gender) and women (regardless of race/ethnicity) performance on tests was significantly lower when faced with a threatening situation compared to White Americans and White American men, respectively. Focusing on these two groups separately does not address the effect of stereotype threat of those who experience multiple and intersecting forms of oppression, such as African American women.

With notable exception (Wright-Adams, 2014), there is a dearth of research examining stereotype threat among African American women. Wright-Adams (2014) examined the effects of stereotype threat among a sample of African American women college students and the results were unexpected. Participants were read either a threatening (diagnostic instructions that mentioned race and gender) or a non-threatening script (did not mention race and gender) from a Black woman experimenter prior to taking a math task. The participants assigned to the threatening condition performed *better* than those assigned to the non-threatening group. Two factors might have

attributed to the unexpected results. First, the African American woman experimenter's presence might have boosted performance among the participants. Steele (2010) mentioned that having Black individuals in power (e.g., Black professor) could alleviate stereotype threat and Nguyen and Ryan (2008) stated that an in-group experimenter might act as a stereotype threat-removal. Second, the participants in Wright-Adams' study completed the math task in a group setting with other African American women, which mirrored a predominantly Black environment. This type of setting could have caused a "critical mass" effect. Critical mass is when negative contingencies are improved because there are a significant number of minorities in a setting (Steele, 2010). Additional factors, such as these, should be assessed in future stereotype threat research.

STEREOTYPE THREAT MODERATORS

There are several known moderators of the effects of stereotype threat on academic performance (see Nguyen & Ryan, 2008 for a list). Some of these moderators include domain identification (Nadler & Komarraju, 2016; Nguyen & Ryan, 2008; Steele, 2010; Walton & Cohen, 2003), stereotype threat type (Nguyen & Ryan, 2008), test difficulty (Nguyen & Ryan, 2008; Steele, 2010; Walton & Cohen, 2003), and racial identity attitudes (Davis et al., 2006; Ho & Sidanius, 2009; McFarland et al., 2003; Wright-Adams, 2014). When considering the pervasiveness and impact of stereotype threat among White women, White men, ethnic minority women, and ethnic minority men, it is difficult to posit that the phenomenon is solely contributing to impaired performance.

Domain identification can be described as how one identifies with a particular domain, or area. In the aforementioned example about women and math performance,

women who moderately cared about their math performance were more susceptible to stereotype threat, which can be described as moderate domain identity (Nguyen & Ryan, 2008). Also noted earlier, the level of test difficulty can also determine how the individual will perform in the face of stereotype threat, which is indicative of its moderating ability.

Racial identity is another potential moderator of the link between stereotype threat and academic performance (Craemer & Orey, 2017; Davis et al., 2006; Ho & Sidanius, 2009; McFarland et al., 2003; Wright-Adams, 2014). Studies have yielded mixed findings. Researchers have found that positive racial identity attitudes have helped performance (Davis et al., 2006). Students with greater, internalized positive racial identity attitudes performed the best in the low threat condition. Other studies have found the opposite, where students with more internalized positive racial identity attitudes performed the worst (Craemer & Orey, 2017; Ho & Sidanius, 2010; McFarland et al., 2003). Because findings are mixed in this area and only one study has included racial identity attitudes with African American women, additional research is needed in this area. The current study only included the Internalization Multiculturalist Inclusive subscale (IMCI; Vandiver et al., 2001) to build upon Davis et al.'s findings regarding positive internalized attitudes serving as a buffer to stereotype threat. Using only the IMCI subscale takes a strengths-based approach in investigating potential racial buffers to threatening conditions.

EXPERIMENTER RACE AS A MODERATOR OF STEREOTYPE THREAT

The race of the experimenter has been associated with increased performance (Marx & Goff, 2005; Wright-Adams, 2014) in stereotype threat studies and scholars have

encouraged additional research in this area (Davis et al., 2006). Although the presence of Black experimenters has caused increased test performance for Black participants on practice verbal GRE items, most stereotype threat studies have not included experimenter race as a moderator, and no study has directly observed this effect with Black women participants. It is important to understand how this potential moderator is connected to Black students' performance in the face of stereotype threat to better understand the role of context on performance with stereotype threat.

Steele's (2010) notion of critical mass and Marx and Goff's (2005) findings illustrate how experimenter race could positively affect performance and act as a buffer to stereotype threat. Steele discussed critical mass as a cue that can ameliorate some negative effects of stereotype threat in multiple settings. He defined critical mass as, "to the point at which there are enough minorities in a setting, like a school or workplace, that individual minorities no longer feel uncomfortable there because they are minorities—in our terms, they no longer feel an interfering level of identity threat" (p.135). Moreover, Steele mentioned that critical mass could be as little as one or two other people and that an authority figure with a shared identity can reduce identity threat. In other words, a Black student can feel less intimidated by a challenging test in the face of stereotype threat if the instructor or test proctor is Black. This notion of critical mass supports Marx and Goff's (2005) findings. Black participants in their study reported higher scores when administered the test by a Black experimenter, and they endorsed more stereotypical thoughts when a White experimenter administered the test.

Strong support for experimenter effects has not been found for Black women participants and stereotype threat, but some researchers alluded to the importance of the

investigation (Davis et al., 2006; Wright-Adams, 2014). Specifically, Wright-Adams (2014) suggested that her presence may have caused increased test performance among Black women participants in the threat condition, but the effects of experimenter race were not controlled for or manipulated in her experiment (no control group or the inclusion of non-Black experimenters). There is room in the literature to assess if Black undergraduate women's math test performance is affected by the race of the experimenter.

BLACK RACIAL IDENTITY

William Cross's (1971) Nigrescence theory is one of the first theories in psychology to describe African American racial identity. In his original model, Cross proposed that psychological Black liberation from racial oppression included stages that African Americans must undergo in order to change from "Negro" (less aware with internalized self-hatred) to "Black" (more aware and committed to change), and understanding what it means to identify as African American is not simple, but rather complex. The stages of Cross's original Nigrescence model included: Pre-encounter (European American worldview), Encounter (experiences a significant event that affects beliefs), Immersion-Emersion (immerses in Black identity with hatred of Whites), and Internalization/Commitment (internalized positive Black identity). Individuals from a myriad of backgrounds, such as low or high socioeconomic statuses, can be categorized in the process of Nigrescence (Cross, 1991; Cross et al., 1995).

In 1991, Cross revised his model of Black identity by highlighting the difference between group identity and personal identity and he revised many of the stages by adding three Pre-Encounter subtypes (Assimilation, Miseducation, and Self-Hatred), two

Immersion-Emersion subtypes (Anti-White and Intense Black Involvement), and three Internalization subtypes (Nationalist, Biculturalist, and Multiculturalist Inclusive). Cross (1991) also noted the fluency or “recycling” of identity development (see Parham, 1989; Helms, 1995), which meant that Black individuals could go through statuses and then return to them later in life due to new challenges and life events.

Each Black identity status, Pre-encounter, Encounter, Immersion-Emersion, and Internalization, is often coupled with emotional responses and psychological buffers or threats. Carter (1991) found that Pre-encounter attitudes were related to increased anxiety, distress, and feelings of inferiority. It is not difficult to imagine the amount of psychological distress or trauma one could face after an experience (i.e., encounter) that “shakes” one’s beliefs about Black identity (e.g., police brutality). In a racial life narrative study, Black individuals who experienced racial epiphanies or “awakenings” became more aware of racism, challenges to fight racism, and strengths of the Black community (Neville & Cross, 2017). As noted earlier, individuals in the Internalization status often have peace, or low psychological distress, as it pertains to their Black identity, which Cross (1991) considered a psychological buffer or defense in racially threatening situations. Due to the existence of negative stereotypes about African Americans’ intellect, numerous predominantly White schools, and an achievement gap between students of color (Black and Latino/a) and White students, researchers have attempted to understand if and how Black identity attitudes act as moderators for academic and testing performance.

Findings indicate racial identity attitudes are related to academic achievement (Cokley, 2014; Chavous et al., 2003; Sellers et al., 1998), hobbies (Dixon, Zhang, &

Conrad, 2009) and psychological well-being among Black individuals (Barrie et al., 2016; Constantine, Donnelly, & Myers, 2002). Highlighting mixed findings, individuals who positively evaluated their racial group have had higher grade point averages than those who viewed their race less positively; and the opposite was found in other cases (see Cokley, 2014 for a review). It is important to note these studies have not included testing performance post stereotype threat cues.

RACIAL IDENTITY BELIEFS AND BEHAVIORS AS MODERATORS OF STEREOTYPE THREAT

Racial identity can be observed as two components: beliefs and behaviors. Assessing racial identity attitudes addresses the individual's *mindset*, while assessing cultural engagement addresses the individual's *actions*. Cross initially highlighted the importance of actions in the process of Nigrescence by including the fifth status of racial identity—Internalization-Commitment status, where individuals turn their appreciation of being Black to action and commitment. Although this status is no longer included in Cross's racial identity model, it is important to assess if one's ideologies transform into actions, and how these actions affect the individual. Both are important in attempting to understand how one's racial identity play or do not play a role in moderating the effects of stereotype threat. This study attempts to assess one's racial identity beliefs with Cross' (1991) IMCI subscale and one's racial identity behaviors with racial identity behavior items, also referred to as cultural engagement items.

Racial identity beliefs. There are mixed findings on the role of affirmative racial identity attitudes on African American college students' academic performance. Some studies found an association with increased test performance in the face of stereotype

threat (Davis et al., 2006; McFarland et al., 2003). Other studies found a decrease in test performance in the face of stereotype threat (Ho & Sidanius, 2010; McFarland et al., 2003), and in some cases, null effects on test performance were revealed (Wright-Adams, 2014). It appears the time in which the racial identity measures was administered plays a role in performance (McFarland et al., 2003). Davis and colleagues (2006), Wright-Adams (2014), and Ho and Sidanius (2010) all administered the racial identity measures before the task. When administered before the task, racial identity is expected to act as a buffer, but the opposite is predicted to occur when the racial identity measure is administered after the task (McFarland et al., 2003). More is needed to explore whether these findings replicate across studies.

To date, only two studies have connected positive racial identity beliefs, or high racial pride and internalization, to increased academic performance among African American college students (Davis et al., 2006; McFarland et al., 2003). When administering seven items from the RIAS (Helms & Parham, 1990), McFarland and colleagues found that greater, positive internalized racial identity beliefs reduced the effect stereotype threat compared to those with lower racial identity beliefs scores. It should be noted that this study only used seven items from the RIAS (subscale not specified), where Davis et al. (2006) included the entire scale. Davis and colleagues primed the students in the high threat group by administering the MEIM and reading a diagnostic script. Only a non-diagnostic script was read to students in the low threat group. Racial identity beliefs moderated performance when the stereotype threat condition was low. That is, students whose racial identity beliefs were more important to their identity (i.e., higher Internalization attitudes) performed the best in the low threat

condition only (Davis et al., 2006). These results highlight how racial identity beliefs may not be a sufficient moderator alone in high threat conditions, but that other race-related moderators might be attributing to test performance as well.

Davis and colleagues' (2006) findings are important to the stereotype threat literature as they provide some evidence for racial identity's moderating effect on stereotype threat among African American students. The key limitation of this research is that the RIAS was used to measure racial identity attitudes, which has been found to have psychometric limitations concerning reliability and validity (Cokley, 2007; Fischer, Tokar, & Serna, 1998; Harkley, McLellan, & Randall, 2002; Tokar & Fischer, 1998; Yanico, Swanson, & Tokar, 1994). A more psychometrically sound version of the racial identity measure is available for use (i.e., CRIS; Vandiver et al., 2001) to measure racial identity attitudes among African American college students.

Stereotype threat research has also shown that Internalization racial identity attitudes have been connected to poorer testing performance (Ho & Sidanius, 2010; McFarland et al., 2003). Using the RIAS, McFarland and colleagues (2003) found that students with greater Internalization status racial identity beliefs performed the worst in the threat condition. Compared to McFarland et al. (2003), Ho and Sidanius (2010) found similar results among African American undergraduates using the Multidimensional Inventory of Black Identity (MIBI; Sellers, Smith, Shelton, Rowley, & Chavous, 1998). When in the threat condition, students who endorsed relatively high public and private regard were most negatively affected by stereotype threat. That is, students who believed society viewed African Americans more positively (public regard) and who felt positively about African Americans (private regard) were more vulnerable to the

impairing effect of stereotype threat. The researchers concluded that those students had “more to lose” than others since they held on to such positive views regarding how others viewed their group. This is also congruent with previous research that showed students with high public regard were more susceptible to depression, stress, and feeling “bothered” by discrimination after indicating perceived discrimination (Sellers & Shelton, 2003; Sellers, Copeland-Linder, Martin, & Lewis, 2006).

In at least one study, racial identity did not moderate the association between stereotype threat and test performance among African American women (Wright-Adams, 2014). Specifically, Wright-Adams (2014) found a non-significant relation between Internalization attitudes and performance on math problems. Wright-Adam’s results were unique as students in the high threat group performed the best—regardless of racial identity attitudes. Two major differences exist that in Wright-Adams’ study might explain the equivocal findings: 1) African American women’s performance was examined instead of both men’s and women’s performance and 2) standardized *math* test questions were used rather than standardized verbal and cognitive test questions that were used in other studies.

The time in which racial identity measures were administered has been related to testing performance (McFarland et al., 2003). Although participants who more positively evaluated their racial groups performed the lowest in the threat condition, those with higher racial identity beliefs performed the best when the scale was administered *before* the cognitive task but those with higher racial identity attitudes performed the worst when they were given the racial identity scale *after* the cognitive task (McFarland et al., 2003). This could possibly explain Ho and Sidanius’ (2010) findings: when the racial identity

measure was given *after* the cognitive task—students with more positive attitudes about being Black performed the worst. It should be noted that although differences were examined in McFarland and colleagues’ (2003) study, they were not significant and when using the pre-post racial identity difference scores, those who endorsed a decrease in race salience performed the best. The researchers asserted that the findings provided some explanation regarding the equivocal stereotype threat and racial identity results. However, they mentioned they trusted the pre-racial identity measure more than the post-racial identity measure, as people could have tried to “dis-identify” from their race after facing difficulties with the cognitive task.

In sum, researchers have examined racial identity beliefs as moderators in stereotype threat studies among African American undergraduates and findings are equivocal; some studies suggest racial identity beliefs are connected to better performance (Davis et al., 2006) and other studies found a connection to poorer performance (Ho & Sidanius, 2010; McFarland et al., 2003). Researchers have also found null effects when measuring racial identity attitudes with math performance among African American women (Wright-Adams, 2014). Moreover, the time in which racial identity measures are administered seems to be important (McFarland et al., 2003), where students with more positive racial identity attitudes performed better when the racial identity measure was given to them before exposure to the threat.

Racial identity behaviors. Participating in culturally-related activities are connected to positive racial identity beliefs (Cokley & Helm, 2007; Cross, 1991) and increased academic achievement (Brannon et al., 2015) among African American undergraduates. Cultural engagement, or one’s participation in culturally related

activities, is associated increased academic performance (Brannon, Markus, & Taylor, 2015). Brannon and colleagues (2015) connected the concept to W.E.B. DuBois' double consciousness (i.e., feeling like one's identity is divided) and an interdependence schema among African American undergraduates. Interdependence was exhibited when participants were more cooperative and creative while completing tasks after they were primed with images from African American culture (Brannon et al., 2015) compared to the control group, that was primed with images from White American culture.

Research findings indicate that engaging in culturally-related activities leads to a more salient independent schema, which positively predict students' grade point averages, academic fit, and academic persistence (Brannon et al., 2015). Cultural engagement also predicts a positive sense of connection to one's racial and ethnic group. Lent and colleagues' (2005) findings support the notion that engagement with one's culture boosts academic performance. In their study, engineering students at two historically Black colleges reported significantly higher academic self-efficacy, outcome expectations, technical interests, social supports, and goals than did students at the predominantly White university (Lent et al.). In sum, cultural engagement has been connected to positive racial identity attitudes and although it has been shown to increase academic achievement, it has not been included in stereotype threat research as a moderator. Racial identity behaviors should be examined as a moderator because they could increase performance, or act as buffer, in the face of stereotype threat.

PURPOSE OF THE PRESENT INVESTIGATION

The purpose of this study was to address the gaps in stereotype threat literature by exploring the potential moderating roles of the race of the experimenter and racial

identity beliefs and behaviors on the link between stereotype threat and math performance among Black women. Taking a strengths-based approach in understanding if and how positive internalized racial identity acts as a buffer to stereotype threat, only one component of racial identity beliefs was measured—Internalization Multiculturalist Inclusive (IMCI; Vandiver et al., 2001). Also, the current study builds on prior research, in which positive internalized racial identity beliefs reduced the effects of stereotype threat compared to less positive racial identity beliefs (Davis et al., 2006; McFarland et al., 2003). There is only one study that examined math performance and stereotype threat among Black women, and there is no study that includes multiple race-related moderators. Some studies have examined race-related moderators such as racial identity or experimenter race separately, but there are no studies that have included them together. This is important because one’s racial identity can be associated with one’s actions, and researchers are unsure how students’ performance is affected by the race and gender of the experimenter. Thus, this study will extend Davis et al.’s (2006) and Wright-Adams’ (2014) research by assessing if the race of experimenters (i.e., White woman or Black woman) moderate test performance with stereotype threat. As an exploratory approach, cultural engagement or racial identity behavior will also be assessed as a potential moderator of test performance.

CHAPTER 3: METHOD

PARTICIPANTS

Participants were 103 undergraduate Black women recruited from a predominantly White university in the Midwest. All participants racially identified as Black. The majority (70.8%; $n = 73$) ethnically identified as African American and the remaining participants identified with other Black ethnic backgrounds. A number of people indicated they had some Nigerian (12.6%; $n = 13$) or other African ancestry (4.9%; $n = 5$); a few identified Caribbean heritage (5.8%; $n = 6$). Four individuals (3.9%) reported Black as their ethnicity and two participants (2%) did not indicate an ethnicity.

The mean age of sample was 19.66 years ($SD = 1.35$; range = 18-25). One third of the sample was freshmen ($n = 35$); another 27 were sophomores (26.2%), 23 were juniors (22.3%), and 18 were seniors (17.5%). Twenty-nine students participated in the study for extra credit in psychology and educational psychology courses. Seventy-four participants who did not choose extra credit received \$5 cash as a token of appreciation for their participation.

DESIGN

The study was a 2 (threat condition) x 2 (experimenter race) between-subjects factorial design. The two independent variables were stereotype threat condition (low threat and high threat) and the race of the experimenter (White and Black). The dependent variable was the participant's score on a math task.

POWER ANALYSIS AND SAMPLE SIZE CALCULATION

In order to determine if the statistical tests would have sufficient power, the required size of N , sample size, was calculated. Using a 90% confidence interval, the calculations yielded a minimum sample size of 102 participants. The total number of participants was divided by four (the number of experimenters). The quotient was 25.5, which I rounded up to 26 participants per experimenter. Thus, the targeted total sample size for the study was 104 (26 participants designated to each of the four experimenters; 13 per threat condition). The following formula was used to calculate the appropriate sample size: $N = \left(\frac{z\sigma}{E}\right)^2$. The formula indicates that the sample size was determined by the squared quotient of the standard deviation z -score divided by the margin of error. The mean ($M = 2.05$), standard deviation ($SD = 1.01$), and error bound for the population mean (EBM or E; $EBM = .21$) were derived from Wright-Adam's (2014) study as the same math task measure was used for the current study. Although using a different mode of calculation, the same sample size was reported in her study as well, using a p -value of .05, effect size of .05 and power of .80 (Creswell, 2003; G.Power3.1).

MEASURES

Demographic information. Participants completed a demographic survey that included questions regarding their age, race, ethnicity, gender, year in school, college, major, and parents' level of education (see Appendix A).

Participants' academic performance. ACT math scores were obtained through the university Registrar's Office after participants granted written permission.

Racial identity beliefs: The Cross Racial Identity Scale (CRIS; Vandiver et al., 2001). The CRIS is a 40-item self-assessment scale grounded in the expanded

Nigrescence model (Cross & Vandiver, 2001). The full measure includes 6 subscales that represent three Black identity types: Pre-Encounter (assimilation, miseducation, and self-hatred), Immersion-Emersion (anti-European American), and Internalization (Afrocentricity; Multiculturalist Inclusive, IMCI). To ensure students would complete the online survey and consistent with previous research, only the Internalization Multiculturalist Inclusive (IMCI) subscale was included in the study. Items are rated on a 7-point Likert-type scale ranging from 1 (*strongly agree*) to 7 (*strongly disagree*). The following is a sample item from the IMCI subscale: “As a multiculturalist, I am connected to many groups (Hispanics, Asian-Americans, Whites, Jews, gays & lesbians, etc.)” (Worrell, Vandiver, & Cross, 2004, p. 14). Vandiver et al. (2002) provided support of the CRIS’s reliability and validity by comparing the measure to different ethnic measures. Among African American college students, CRIS subscale alpha coefficients have ranged from .69 (Pre-Encounter self-hatred; Wright-Adams, 2014) to .89 (Immersion-Emersion anti-European American; Vandiver et al., 2002). The alpha coefficients range from .69 (Pre-Encounter self-hatred; Wright-Adams, 2014) to .87 (Pre-Encounter assimilation; Wright-Adams, 2014) for African American women alone. The alpha coefficient for the IMCI subscale in the current study was .81.

Racial identity behavior. A modified version of Brannon et al. (2015) indicator of racial identity behavior or cultural engagement was used. In their study, Brannon and her colleagues assessed the construct with two items: “I participate in events or activities sponsored by groups reflecting my own cultural heritage” and “I rely on racial/cultural groups as my main support group on campus” (p. 600). I added four additional items to increase the reliability of the scale. On the basis of Brannon et al.’s definition of cultural

engagement as “respondents’ engagement with African American culture” (p. 600), the following items were constructed: “I am an active member of an African American organizational club/group,” “I attend lectures/talks on campus about African Americans,” “I consume information about African Americans via social media, books, articles, etc.,” and “I share information related to African Americans with others via social media, text messages, and/or conversations” (see Appendix B). The items were finalized in consultation with Brannon (2016). Responses are rated on a 4-point Likert-type scale, from 1 (*strongly disagree*) to 4 (*strongly agree*). Brannon et al.’s indicator of cultural engagement was linked to increased academic performance and persistence among a sample of African American college students. The alpha coefficient for the racial identity beliefs scale was .82.

Math performance: Practice SAT math items. Wright-Adams (2014) randomly selected math items from *The Official SAT Study Guide*, 2nd edition (The College Board and Educational Testing Service, 2009). Nine of these items were included in the study to assess math performance. A total sum score was calculated for each participant.

PROCEDURE

Recruitment. A secure electronic list of all African American women was acquired from the campus Division of Management Information (DMI). All participants were sent an email that included a link to access the online portion of the study. Participants were initially told that the study consisted of two parts. Part one: participants were asked to complete an online survey consisting of the demographic items, cultural engagement questions (racial identity behavior), the CRIS IMCI subscale (racial identity beliefs), and scheduling options to meet in person for the second portion of the study. The

online consent form also included an explanation of part two, but the stereotype threat manipulation was not mentioned. Specifically, participants were informed:

In Part Two, you will be asked to spend 30 minutes or less answering math questions. You will then be debriefed about the study and will be asked to sign a consent form to give the primary researcher the right to obtain your ACT/SAT scores from the University Office of Registrar.

For Part Two of the study, participants completed the math task in a lab setting. After scheduling a date and time to complete Part Two, participants were randomly assigned to the low or high threat group and a Black or White woman experimenter.

Experiment. For Part One of the study, participants were asked to complete the online survey (includes aforementioned measures) and to schedule an appointment to complete Part Two, which was at least seven days after they completed part one. The participants were randomly assigned to the low or high threat group once the appointment for Part Two was confirmed. These assignments also included random assignment with a White or Black woman experimenter.

In Part Two, the participants were presented with 10 math problems individually: “Solve each problem and decide which is the best of the choices given by circling the best response. You may use any available space for scratch work.” Ten math items were administered but an error was found in a response option for one item. This left a total of nine items for statistical analyses. In the low threat group, participants were told, “This research is aimed at better understanding what makes some people better at math than others” (Wright-Adams, 2014; see Appendix D for full script). Conversely, in the high threat group, participants were told, “A good deal of research indicates that White and

Asian men consistently score higher than Blacks, especially Black women on standardized tests of math. Research also shows that White and Asian women perform better than Black women. But thus far, there is not a good explanation for this” (see Appendix E for full script). Therefore, the experimental conditions consisted of (1) low threat: no race and gender prime with problem-solving instructions and (2) high threat: race and gender prime with mathematical intelligence instructions.

Participants were given 25 minutes to complete the math task after instructions were given. Prior to the math task, students were asked to complete math identification items, and were then asked to complete difficulty, effort, and performance items after the math task. Participants were debriefed about the study after completing the project related tasks and were asked to sign a waiver to give the primary researcher the right to obtain their ACT math scores from the registrar’s office for data analysis. Six participants declined the right of the researcher to obtain their ACT records and two participants did not have any ACT scores on file. They were given \$5 dollars cash or extra credit for psychology or educational psychology courses, a resources list for Black women, and a snack as tokens of appreciation.

Training. The experimenters received two hours of training before data collection. The training included an orientation to the topic and the research, a review of the protocol, scripts, and dress code (casual attire was prohibited). All experimenters were given a binder with copies of the instructions on how to conduct the study, a checklist, and the necessary materials (see Appendices G and H). All participants completed the math tasks in a lab room with the designated experimenter. Experimenters were instructed to adhering to the protocol to minimize extraneous experimenter effects.

STATISTICAL ANALYSES PLAN

Preliminary analyses. Pearson product-moment correlations were calculated to explore if math ACT scores were related to the study variables. Pearson product-moment correlation was also conducted to examine associations between all quantitative variables (see Appendix J). ACT math scores were significantly related to math task performance, math course enrollment, math identification, math task difficulties, and perceived math task performance.

Main analyses. A 2 x 2 Analysis of Covariance (ANCOVA) was performed to test the first two research questions: (a) if participants in the low stereotype threat condition performed better on the math task than participants in the high stereotype threat condition, and (b) if there was an interaction between experimenter race and condition. Last, to examine if racial identity beliefs and racial identity behaviors moderated the association between stereotype threat and math performance, I conducted a hierarchical multiple regression on math task performance. Predictor and moderator variables were centered prior to the analysis. First, ACT math scores were entered into the first step. Second, stereotype threat condition, experimenter race, racial identity beliefs, and racial identity behaviors were entered. Third, interactions between stereotype threat and beliefs and stereotype threat and behaviors were entered.

CHAPTER 4: RESULTS

DATA CLEANING

Univariate and multivariate outliers, skewness, and kurtosis were examined using IBM SPSS software (Version 24.0). Univariate outliers were identified through boxplots and histograms, and outliers were transformed. Four outliers were identified for the CRIS IMCI subscale and were adjusted by transforming outlier values to the next lowest (non-outlier) value. Outlier analyses were conducted again and no outliers were found after the transformation.

MISSING DATA

Missing data patterns were assessed using Little's *missing completely at random* (MCAR) test, and the null hypothesis was not rejected (Little & Rubin, 2002). This indicated that the data were missing at random, $\chi^2(52, N = 103) = 51.28, p = .50$. The dataset contained a relatively small percentage of missing values (2.09%). Expectation Maximization (EM) was used to input missing data and this method has been proven to obtain acceptable reliability estimations with a small percentage of missing data (Cuesta & Fonseca, 2014).

PRELIMINARY ANALYSES

Sample. To assess if the participants who chose cash compensation were comparable to participants who received course credit, an independent-samples *t* test on racial identity beliefs was conducted to compare students based on compensation choice (cash v. credit). The *t* test revealed no significant differences between the participants that chose cash as a form of compensation and those who chose extra credit towards psychology or educational psychology courses on the IMCI subscale, $t_{(101)} = 1.52; p >$

.05. A second independent-samples t test on math task score grouped by compensation choice was performed. Again, no significant differences between the groups were found, $t_{(101)} = 1.39$; $p > .05$. Due to the aforementioned results, the two groups were treated as one sample for the remaining analyses.

ACT math scores. Two ACT math score variables were examined in the preliminary analyses: (a) the first ACT variable included the ACT scores of all participants who granted the researcher permission to obtain scores that were on file with the registrar's office ($N = 95$, $M = 23.52$, $SD = 3.63$). This variable included eight missing ACT values (six participants declined and two did not have any scores on file), and (b) the second ACT variable included an ACT math score for all participants ($N = 103$, $M = 23.55$, $SD = 3.48$). The eight participants who did not have scores were given the rounded mean score, 24, from the sample, as this method was followed in Davis et al.'s (2006) study. The minimum ACT score was 16 ($n = 3$) and the maximum was 33 ($n = 2$). The mode was 24 ($n = 26$). Fisher's exact test revealed that there was not a significant relationship between granted ACT access and experimenter race ($p = 1.00$). Meaning, the experimenter's race was not connected to the participant's willingness to grant access to their ACT math scores.

To ensure that giving the eight participants the sample mean ACT score did not significantly impact the results, two additional missing data procedures were performed—EM and listwise deletion. Two copies of the original dataset with missing ACT scores were made, and the EM procedure was performed on one dataset and the listwise deletion was conducted on the other dataset. Imputed values from the EM procedure were analyzed and it was found that the EM algorithm imputed the sample

mean ACT score (24) for each missing value. Because the values did not differ from the values that I initially imputed based on the method used in Davis et al. (2006) research, the main analyses were only conducted again for the listwise deletion method for comparison purposes. Findings from the main analyses did not indicate significant changes between the listwise deletion or mean imputed datasets, therefore, the method that assigned the sample mean ACT score was retained. In other words, the eight participants who did not grant access to obtain their ACT scores and who did not have any scores on file, were assigned a 24 ACT math score for all of the analyses. Refer to Appendix K for listwise deletion findings.

Assumptions for ANCOVA. Assumptions for ANCOVA were assessed. The first assumption is that values of the covariate (ACT scores) cannot vary across the different levels of the independent variable (threat condition). This was tested by conducting a one-way ANOVA, where ACT scores were entered as the dependent variable and stereotype threat condition was entered as the fixed factor. For comparison purposes, this procedure was completed on the two ACT variables (one missing eight ACT scores and the no missing ACT scores) in two separate analyses. The results were not significant for either statistical test; ACT variable with missing scores, $F(1, 94) = 2.80, p = .10$; ACT variable with no missing scores, $F(1, 102) = 2.89, p = .09$.

The second assumption was assessed—homogeneity of regression. The assessment of this assumption was done to examine if the slopes for threat condition and ACT scores were homogenous. ANOVAs were conducted as above for the dataset with missing ACT scores and the dataset with non-missing ACT scores, but the interaction between ACT scores and threat condition were assessed. The results for ACT variable

without missing scores were not significant, $F(1, 102) = 3.65, p = .06$, thus not rejecting the null hypothesis. The results were similar for the interaction between condition and ACT variable with missing values, $F(1, 94) = 3.53, p = .06$. These findings indicated that ACT math scores served as an acceptable covariate and that there was no significant difference between the two ACT variables (with missing and no missing values) being used as the covariate. ACT scores with imputed means for missing values were used throughout the analyses to maximize statistical power.

ACT math scores were related to math task but not racial identity. The Pearson product-moment correlation revealed a significant positive association between ACT scores and math task score, $r(101) = .29, p = .003$. Students with higher ACT scores scored higher on the math task. There were no significant relationships between ACT scores and racial identity beliefs ($r[100] = .05, p = .64$) and racial identity behaviors ($r[100] = .07, p = .49$). See Appendix J for associations for all continuous variables. Refer to Appendix L for additional variables' associations with ACT math scores.

EFFECTS OF STEREOTYPE THREAT ON MATH PERFORMANCE

To investigate if math performance was affected by low threat and high threat conditions after controlling for prior math performance, a 2 X 2 ANCOVA was conducted. Only observations in which I received official ACT math scores were included in the analysis. The high threat condition included a race and gender prime with mathematical intelligence instructions and the low threat condition did not include a race and gender prime, and it included problem-solving instructions only. The results did not reveal a main effect by threat condition after controlling for ACT math scores, $F(1, 102)$

= 1.20, $p = .28$. See Table 1 for descriptive statistics of conditions. The results were similar for African American participants only, $F(1, 71) = .60, p = .44$.

Table 1: Mean score on math task per group.

Condition	<i>M</i>	<i>SD</i>	N
Low threat	3.20	1.30	51
High threat	3.33	1.15	52

Note. Mean math score was the average score on the math task (nine math items). The mean difference between low threat and high threat groups was not statistically significant.

EFFECTS OF EXPERIMENTER RACE ON MATH PERFORMANCE

The results from the previous 2 X 2 ANCOVA were also used to address the second research question regarding the effects of experimenter race on math performance. The main effect for experimenter race was not significant, $F(1, 102) = 1.11, p = .30$. Moreover, a main effect for the experimenter race and stereotype threat condition interaction was not significant $F(1, 102) = .57, p = .45$. In other words, participants' performance did not significantly differ based on threat condition or experimenter race assignment in any way. Refer to Table 2 for descriptive statistics.

To ensure that there were no extraneous experimenter effects, a separate ANCOVA was conducted with math task as the dependent variable, experimenter assignment (i.e., experimenter 1, 2, 3, or 4) as the independent variable, and ACT scores as the covariate. There was no significant main effect found for experimenters by assignment, $F(1, 102) = 1.79, p = .15$; thus, ruling out personality or other potential factors.

Table 2: Mean score on math task by threat condition and experimenter race.

Race of Experimenter	Threat Condition	Mean	SD	N
Black	Low	3.38	1.27	26
	High	3.54	.99	26
	Total	3.46	1.13	52
White	Low	3.00	1.32	25
	High	3.12	1.28	26
	Total	3.06	1.29	51

Note. Means represent the average scores on the math task for participants randomly assigned to White or Black experimenters and high or low threat conditions. The mean difference between those who were assigned to a Black or White experimenter was not significant ($p = .30$). The interaction between experimenter race and threat condition was also not statistically significant ($p = .45$).

RACIAL IDENTITY BELIEFS AND BEHAVIORS RELATIONSHIP WITH MATH PERFORMANCE

To examine if racial identity beliefs and behaviors moderated the effects of stereotype threat on a math task among Black women, a multiple regression was conducted. Pearson product-moment correlation was used to examine if there was an association between ACT scores and racial identity beliefs (IMCI subscale) and racial identity behaviors (cultural engagement items). The results did not indicate significant relationships between ACT scores and the two variables: racial identity beliefs ($r[100] = .05, p = .64$) and behaviors ($r[100] = .07, p = .49$). There were no significant correlations between the independent variables ($r[100] = .18, p = .07$) and the variance inflation factor was below 1.2, which indicated that the data were free from collinearity issues (Chatterjee & Price, 1991; Draper & Smith, 1981). Pearson product-moment correlation indicated there were not significant associations between math task score and racial identity beliefs ($r[100] = .18, p = .07$) and racial identity behaviors ($r[100] = .19, p = .19$).

To assess if racial identity (beliefs and behaviors) moderated the effects of stereotype threat on math performance, a multiple regression was conducted. The participants' composite math scores were entered as the dependent variable. For the first model, threat condition was dummy coded 0 (low threat) and 1 (high threat) and entered as a fixed factor. The ACT math scores were entered in the model as the covariate. For the second model, racial identity beliefs mean scores and racial identity behaviors mean scores were added to the model. For the third model, all variables were retained and the interaction terms, threat condition x racial identity beliefs and threat condition x racial identity behaviors, were added. All variables were centered prior to analysis.

The results from the hierarchical multiple regression revealed that the combination of predictors significantly predicted math task performance for each model. Model 1 (threat condition and centered ACT scores) significantly predicted math task performance, $F(2, 99) = 4.68, p = .01$. Model 2 (threat condition, centered ACT scores, centered racial identity beliefs and behaviors) also significantly explained math task performance $F(4, 97) = 4.23, p = .003$. Last, Model 3 (threat condition, centered ACT scores, centered racial identity beliefs and behaviors, and threat condition and racial beliefs interaction, and threat condition and racial identity behaviors interaction) was also significant in predicting math task scores $F(6, 95) = 3.42, p = .004$. It should be noted that ACT math scores accounted for most of the variance within the models, which could greatly contribute to the p-value of each model. Recall, the other predictors were not correlated with math task.

Racial identity beliefs and racial identity behaviors did not contribute significantly to the regression models (see Table 4). Racial identity behaviors indicated a significant

effect in the third model, but this should be interpreted with caution, as the variable was not significant in the second model where the stereotype threat x racial identity behaviors interaction was not included. Moreover, none of the interaction terms significantly predicted math task performance. The model with all predictors (model 3) accounted for approximately 13% of the variance in math task scores (see Table 3). ACT math scores showed to be a significant predictor in every model.

Table 3: Model summary for multiple regression analysis.

<i>Model</i>	R	R Square	Adjusted R Square
1 ^a	.29	.09	.07
2 ^b	.39	.15	.11
3 ^c	.42	.18	.13

Note. Adjusted R Square indicates the amount of variance that is accounted for in math task scores per model. Model 1 predictors: threat condition and centered ACT scores. Model 2 predictors: threat condition, centered ACT scores, centered racial identity beliefs and behaviors. Model 3 predictors: threat condition, centered ACT scores, centered racial identity beliefs and behaviors, and threat condition and racial beliefs interaction, and threat condition and racial identity behaviors interaction. All models were significant all the .05 level.

Table 4: Multiple regression analysis for variables predicting math task performance ($N = 102$).

<i>Variable^a</i>	B	SE B	β	<i>p</i>
Model 1				
ACT	.10	.03	.30**	.00
Stereotype threat	.21	.23	.09	.36
Model 2				
ACT	.10	.03	.29**	.00
Stereotype threat	.36	.24	.15	.13
RI Beliefs	.26	.15	.17	.08
RI Behaviors	.20	.12	.16	.10
Model 3				
ACT	.09	.03	.26**	.01
Stereotype threat	.36	.23	.15	.12
RI Beliefs	.27	.22	.18	.23
RI Behaviors	.43	.17	.34**	.01
Ster. Threat x RI Beliefs	.05	.30	.02	.87
Ster. Threat x RI Behaviors	-.44	.24	-.26	.07

Note. Significant at the 0.01 level (2-tailed).**

^a Covariate (ACT) and all independent variables centered on mean. Stereotype threat condition was dummy-coded (0—low threat; 1—high threat). Racial identity beliefs (RI Beliefs) were measured with CRIS IMCI subscale. Racial identity behaviors (RI Behaviors) were measured with cultural engagement items.

CHAPTER 5: DISCUSSION

The purpose of this study was to examine the effects of stereotype threat on math performance among Black undergraduate women at a predominantly White institution. Stereotype threat phenomenon was first examined on a PWI campus (Steele & Aronson, 1995), and this setting has also been included in other stereotype threat studies (e.g., Cromley et al., 2013; Wright-Adams, 2014). In general, there were no statistically significant differences on the math task between participants who were randomly assigned to high or low threat conditions after controlling for ACT math scores. Inconsistent with previous research, the use of moderators, such as experimenter race and racial identity, did not significantly improve performance on math task in either threat condition (Marx & Goff, 2005; Davis et al., 2006). In this study, ACT math scores were the best predictors of math task performance. This is intuitive because the math task was derived from a SAT practice test, which is a standardized test like the ACT.

DID STEREOTYPE THREAT MATTER?

Similar to some stereotype threat research (Cromley et al., 2013; McFarland, Lev-Arey, & Ziegert, 2003), significant mean score differences between low threat and high threat groups were not found on the math task. This means that the Black women participants' math performance was not hindered by highly threatening conditions, which can allude to resiliency or an unpredicted stereotype threat removal.

When also studying Black women's math performance with stereotype threat, Wright-Adams (2014) found that the high threat group performed the best on the math task, and this performance was significantly better than the low threat group. The current study used the same math task and very similar threat scripts, yet statistical significance

was not reached in the current investigation as it was in Wright-Adams' study. It should be noted that Wright-Adams tested participants in a group setting, while the current study tested participants individually. Testing in a group setting could have created a performance boost due to the setting being predominantly Black and producing a critical mass effect.

It could also be interpreted that the use of women experimenters acted as a stereotype threat removal, thus allowing participants in the high threat to perform just as well or slightly better than those with minimal threat. This builds upon Nguyen and Ryan's (2008) meta-analytic conclusions that highlighted how experimenters from the in-group (i.e., shared identity) could act as a stereotype threat removal. Additional research has supported a gender in-group effect (Marx & Roman, 2002). When assessing the effects of experimenter gender and stereotype threat, the presence of women experimenters increased math performance for women participants, while men experimenters caused a decrease in performance (Marx & Roman, 2002). In the case of this study, the use of women experimenters (regardless of race) might have lessened the effect of stereotype threat for Black women participants and created a more "neutral" ground for testing, irrespective of threat condition.

DID EXPERIMENTER RACE MATTER?

Unlike previous research regarding experimenter race and stereotype threat (Marx & Goff, 2005), scores on the math task did not significantly differ by the race of the experimenter. In other words, Black women participants did not significantly perform better or worse when administered the math task by a Black woman or White woman.

The null findings and gaps between the current investigation and Marx and Goff's (2005) could be attributed to participants' demographics and the type of threat. Marx and Goff's (2005) sample was mixed gender (although the number of men and women was not given), while the current experiment included Black women only. It could be posited that Black men might respond differently to women experimenters in stereotype threat conditions. Moreover, the type of threat that was tested significantly differed between the two studies; Marx and Goff examined *verbal* ability while this study examined *math* ability. Verbal ability is a common anti-intellectual stereotype for all Black individuals, regardless of gender, while math ability is usually a gender-specific stereotype that targets women (Spencer et al., 1999; Steele & Aronson, 1995).

Another interpretation of the findings could be related to participants' assumptions about the experimenters' math competence. Undergraduate women participants (race not reported) performed better on a math task when they perceived the experimenter to be competent in math than participants who perceived their woman experimenter to be less competent in math (Marx & Roman, 2002). The women participants in the present study could have assumed that the women experimenters were *all* knowledgeable about math, regardless of race, thus minimizing stereotype threat effects and yielding null results. This was supported by a few participants asking experimenters for clarity on the math items while testing.

DID RACIAL IDENTITY MATTER?

Racial identity beliefs and behaviors did not significantly predict math task scores in the high or low threat condition. Regarding racial identity *beliefs*, the findings deviated from previous research that included both Black men and women (Davis et al., 2006), but

findings were consistent with the aforementioned study that included Black women only (Wright-Adams, 2014). When studying men and women participants, Davis and colleagues (2006) found that high Internalization attitudes predicted verbal task performance in the low threat condition. Examining Black women only, Wright-Adams (2014) revealed null findings between Internalization attitudes and math task performance. Put another way, internalized positive racial identity beliefs acted as a buffer in the low threat condition for Black men and women when completing a verbal task, but this did not hold true for Black women completing a math task.

One should also consider that racial identity beliefs may not be a sufficient protective factor in threatening conditions for those with multiple marginalized identities, such as Black women. Davis and colleagues (2006) mentioned that it is possible that strong race-based attitudes do not act as buffers from all psychological threat, which limits their effects on task performance in the face of stereotype threat. Racial-gendered identity may serve as a better moderator of the stereotype threat - math performance association. A racial-gendered identity measure would assess how one feels about their racial-gendered group, such as Black women instead of Black people or women alone.

Regarding racial identity *behaviors*, Brannon and colleagues (2013) found that cultural engagement (racial identity behavior) was significantly and positively connected to increased academic achievement and persistence, which was contrary to the nonsignificant results in the present study. While the differences between the results could be attributed to the racial-gendered perspective for racial identity attitudes addressed above, the dependent measures could have also made a difference. Brannon and colleagues used GPA as an outcome variable and I used practice math SAT items.

The notion of racial identity behaviors is exploratory in stereotype threat research and the dearth of studies limits comparison of findings. Additional research is needed to better understand findings.

LIMITATIONS AND FUTURE RESEARCH

There are a number of noteworthy limitations in the present investigation. One includes not having Black and White men experimenters to further examine the effects of race and gender. Including men experimenters would have allowed comparisons between gender *and* race within the study. Although the current study used a strengths-based approach and built upon existing literature by only including the IMCI subscale, only using one subscale from the CRIS limited the interpretation of results, as other research have found some excluded subscales to be related to task performance with stereotype threat. Another limitation of the study centers around the issues of replicability and stereotype threat. Scholars have critiqued the stereotype threat literature by highlighting inconsistent findings (Nguyen & Ryan, 2008; Wax, 2009). This can be attributed to different methods in invoking stereotype threat, various outcome variables, many moderators, and multiple ways of measuring the effects of the stereotype threat.

Future research should examine stereotype threat conditions that include Black and White men experimenters in addition to Black and White women experimenters. Because the findings indicated that past performance (ACT math scores) and course enrollment was related to performance, future research could also include measures or qualitative interviews about how participants prepare for math tests and how they perceive classroom environments when testing. Furthermore, future projects related to Black undergraduate women and stereotype threat should incorporate qualitative

methods, such as focus groups. This might help explain how Black women undergraduates perceive racial and gender threats, which could help researchers understand why the results in the current and Wright-Adams' (2014) study were not consistent with most stereotype threat research. Future research should strive to keep stereotype variables, methods, and settings (individual or group) consistent, especially when comparing results in stereotype threat studies.

IMPLICATIONS OF FINDINGS

Unexpected findings highlight how studying Black women and stereotype threat is unique to the literature. Because there are not several stereotype threat studies that utilize Black women and women experimenters from other racial backgrounds, there are limitations in comparing the results. Replications and extensions are necessary to further explore if and how stereotype threat plays a role with this specific population and math performance. Studying other women of color who face anti-intellectual stereotypes, such as Latina women, is also encouraged. There is a dearth of stereotype threat research that considers the intersection of race and gender (i.e., Wright-Adams, 2014). It is important for researchers to use an intersectional approach throughout their stereotype threat work—theory, conceptualization, and reporting participants' demographics (including both race and gender). First, researchers can further examine the strengths and issues associated with identifying as a Black woman. Moreover, incorporate frameworks, such as intersectionality, that address both visible identities in the conceptualization process. Last, stereotype threat researchers should thoroughly explain how participants in the study identified. Only including the participant's race or gender should not suffice.

The results from the current study can inform educational practices with Black women undergraduates. Firstly, schools should hire more women teachers for math courses. The findings revealed that even after presented with a threat, the students' performance was not harmed. This could potentially be attributed to inclusion of women experimenters and the perception of threat due to their presence. Second, because higher ACT math scores were linked to higher math performance, secondary educational facilities should prioritize standardized test prep for Black girls; this might serve as a protective factor in the face of future threatening situations.

REFERENCES

- Aronson, J., Lustina, M., Good, C., Keough, K., Steele, C., & Brown, J. (1999). When White men can't do math: Necessary and sufficient factors in stereotype threat. *Journal of Experimental Social Psychology, 35*, 29–46.
- Barrie, R. E., Langrehr, K., Jérémie-Brink, G., Alder, N., Hewitt, A., & Thomas, A. (2016). Stereotypical beliefs and psychological well-being of African American adolescent girls: Collective self-esteem as a moderator. *Counseling Psychology Quarterly, 29*(4), 423-442.
- Brannon, T. N., Markus, H. R., & Taylor, V. J. (2015). “Two souls, two thoughts,” two self-schemas: Double consciousness can have positive academic consequences for African Americans. *Journal of Personality and Social Psychology, 108*(4), 586-609.
- Carter, R. T. (1991). Racial identity attitudes and psychological functioning. *Journal of Multicultural Counseling and Development, 19*(3), 105-114.
- Chavous, T. M., Bernat, D. H., Schmeelk-Cone, K., Caldwell, C. H., Kohn-Wood, L., & Zimmerman, M. A. (2003). Racial identity and academic attainment among African American adolescents. *Child development, 74*(4), 1076-1090.
- Chatterjee, S., & Price, B. (1991). *Regression analysis by example* (2nd ed.). New York: John Wiley & Sons.
- Cohen, G. L., Garcia, J., Apfel, N., & Master, A. (2006). Reducing the racial achievement gap: A Social-psychological intervention. *Science, 313*, 1307-1310.

- Cokley, K. (2007). Critical issues in the measurement of ethnic and racial identity: A referendum on the state of the field. *Journal of Counseling Psychology, 54*(3), 224-234.
- Cokley, K. O. (2014). *The myth of Black anti-intellectualism: A true psychology of African American students*. Santa Barbara, CA: ABC-CLIO.
- Cokley, K., & Helm, K. (2007). The relationship between African American enculturation and racial identity. *Journal of Multicultural Counseling and Development, 35*(3), 142-153.
- Constantine, M. G., Donnelly, P. C., & Myers, L. J. (2002). Collective self-esteem and Africultural coping styles in African American adolescents. *Journal of Black Studies, 32*(6), 698-710.
- Craemer, T., & Orey, D. A. (2017). Implicit Black identification and stereotype threat among African American students. *Social science research, 65*, 163-180.
- Creswell, J. W. (2003). *Research design: Qualitative, quantitative and mixed methods approaches*, (2nd ed.). Thousand Oaks, CA: Sage Publications, Inc.
- Cromley, J. G., Perez, T., Wills, T. W., Tanaka, J. C., Horvat, E. M., & Agbenyega, E. T. (2013). Changes in race and sex stereotype threat among diverse STEM students: Relation to grades and retention in the majors. *Contemporary Educational Psychology, 38*(3), 247-258.
- Cross, W. E., Jr. (1971). The Negro-to-Black conversion experience. *Black World, 20* (9), 13-27.

- Cross, W. E., Ponterotto, J. S., Casa, J. M., Suzuki, L. A., & Alexander, C. M. (Eds.) (1995). *Handbook of multicultural counseling*. Newbury Park, CA: Sage Press.
- Cross, W. E., Jr., & Vandiver, B. J. (2001). Nigrescence theory and measurement: Introducing the Cross Racial Identity Scale (CRIS). In J. G. Ponterotto, J. M. Casas, L. A. Suzuki, & C. M. Alexander (Eds.), *Handbook of multicultural counseling* (2nd ed., pp. 371–393). Thousand Oaks, CA: Sage.
- Cuesta Izquierdo, M., & Fonseca Pedrero, E. (2014). Estimating the reliability coefficient of tests in presence of missing values. *Psicothema*, 26(4).
- Draper, N. R., & Smith, H. (1981). *Applied regression analysis* (2nd ed.). New York: John Wiley & Sons.
- Davis III, C., Aronson J., Salinas M. (2006). Shades of threat: Racial identity as a moderator of stereotype threat. *Journal of Black Psychology*, 32,399-417.
- Dixon, T. L., Zhang, Y., & Conrad, K. (2009). Self-esteem, misogyny and Afrocentricity: An examination of the relationship between rap music consumption and African American perceptions. *Group Processes & Intergroup Relations*, 12(3), 345-360.
- Field, A. (2000). *Discovering statistics using SPSS for windows*. London-Thousand Oaks- New Delhi: Sage publications.
- Field, A. (2009). *Discovering statistics using SPSS*. London: SAGE.
- Field, A. P. (2013). *Discovering statistics using SPSS: and sex and drugs and rock 'n' roll (4th Edition)*. London: Sage.

- Fischer, A. R., Tokar, D., & Serna, G. S. (1998). Validity and construct contamination of the Racial Identity Attitude Scale—Long Form. *Journal of Counseling Psychology, 45*, 212–224.
- Good, C., Rattan, A., & Dweck, C. S. (2012). Why do women opt out? Sense of belonging and women's representation in mathematics. *Journal of Personality and Social Psychology, 102*(4), 700-717.
- Gravetter, F., & Wallnau, L. (2014). *Essentials of statistics for the behavioral sciences* (8th ed.). Belmont, CA: Wadsworth.
- Hare, B. (2014, October 6). Raven-Symone: I'm not gay, and I'm not African-American. *CNN*. Retrieved from <http://www.cnn.com/2014/10/06/showbiz/raven-symone-gay-labels/>
- Harkley, A., McLellan, A., & Randall, M. (2002). Analysis of the Black Racial Identity Attitude Scale in a substance-abusing population. *Journal of the National Medical Association, 94*, 459–476.
- Heather J. Johnson , Lucy Barnard-Brak , Terrill F. Saxon & Megan K. Johnson (2012). An experimental study of the effects of stereotype threat and stereotype lift on men and women's performance in mathematics. *The Journal of Experimental Education, 80*(2), 137-149.
- Helms, J. E. (1990). *Black and White racial identity: Theory, research, and practice*. New York: Greenwood.
- Helms, J. E. (1995). An update of Helms' White and people of color racial identity models. In J. G. Ponterotto, J. M. Casas, L. A. Suzuki, & C. M. Alexander

- (Eds.), *Handbook of multicultural counseling* (pp. 181-198). Thousand Oaks, CA: Sage.
- Helms, J. E., & Parham, T. A. (1990). *The Black racial identity attitude scale*. New York: Greenwood.
- Ho, A. K., & Sidanius, J. (2010). Preserving positive identities: Public and private regard for one's ingroup and susceptibility to stereotype threat. *Group Processes & Intergroup Relations, 13*(1), 55-67.
- IBM Corp. Released 2016. IBM SPSS Statistics for Windows, Version 24.0. Armonk, NY: IBM Corp.
- Keller, J. (2002). Blatant stereotype threat and women's math performance: Self-handicapping as a strategic means to cope with obtrusive negative performance expectations. *Sex Roles, 47*(3), 193-198.
- Lent, R. W., Brown, S. D., Sheu, H., Schmidt, J., Brenner, B. R., Gloster, C. S., Wilkins, G., Schmidt, L.C., Lyons, H., & Treistman, D. (2005). Social cognitive predictors of academic interests and goals in engineering: Utility for women and students at historically black universities. *Journal of Counseling Psychology, 52*(1), 84-92.
- Marx, D. M., & Goff, P. A. (2005). Clearing the air: The effect of experimenter race on target's test performance and subjective experience. *British Journal of Social Psychology, 44*(4), 645-657.
- McFarland, L. A., Lev-Arey, D. M., & Ziegert, J. C. (2003). An Examination of Stereotype Threat in a Motivational Context. *Human Performance, 16*(3), 181-205.

- Mendoza-Denton, R., Downey, G., Purdie, V. J., Davis, A., & Pietrzak, J. (2002). Sensitivity to status-based rejection: Implications for African American students' college experience. *Journal of Personality and Social Psychology, 83*, 896–918.
- Murphy, M. C., Steele, C. M., & Gross, J. J. (2007). Signaling threat: How situational cues affect women in math, science, and engineering settings. *Psychological Science, 18*(10), 879-885.
- Muthen, L. K., & Muthen, B. O. (2013). Mplus (Version 7.11) Los Angeles. CA: Muthén & Muthén.
- Nadler, D. R., & Komarraju, M. (2016). Negating stereotype threat: Autonomy support and academic identification boost performance of African American college students. *Journal of College Student Development, 57*(6), 667-679.
- Neville, H. A., & Cross, W. E., Jr. (2017). Racial awakening: Epiphanies and encounters in black racial identity. *Cultural Diversity and Ethnic Minority Psychology, 23*(1), 102-108.
- Nguyen, H. D., & Ryan, A. M. (2008). Does stereotype threat affect test performance of minorities and women? A meta-analysis of experimental evidence. *Journal of Applied Psychology, 93*(6), 1314–1334.
- Oliver, A., Andemeskel, G., King, C. R., Wallace, L., McDougal, S., Monteiro, K. P., & Ben-Zeev, A. (2017). ‘I’m black and I’m proud’: A majority ecological context protects affective aspects of black identity under stereotype threat. *Race and Social Problems, 9*(4), 313-320.
- Owens, J., & Massey, D. S. (2011). Stereotype threat and college academic performance: A latent variables approach. *Social Science Research, 40*(1), 150-166.

- Picho, K., Rodriguez, A., & Finnie, L. (2013). Exploring the moderating role of context on the mathematics performance of females under stereotype threat: A meta-analysis. *The Journal of social psychology, 153*(3), 299-333.
- Parham, T. A. (1989). Cycles of psychological nigrescence. *The Counseling Psychologist, 17*(2), 187-226.
- Rivera, L. M., & Benitez, S. (2016). The roles of in-group exemplars and ethnic-racial identification in self-stereotyping. *Social Cognition, 34*(6), 604-623.
- Sellers, R. M., Chavous, T. M., & Cooke, D. Y. (1998). Racial ideology and racial centrality as predictors of African American college students' academic performance. *Journal of Black Psychology, 24*(1), 8-27.
- Sellers, R. M., Copeland-Linder, N., Martin, P., & Lewis, R. (2006). Racial identity matters: The relationship between racial discrimination and psychological functioning in African American adolescents. *Journal of Research on Adolescence, 16*(2), 187-216.
- Sellers, R. M., Smith, M. A., Shelton, J. N., Rowley, S. A. J., & Chavous, T. M. (1998). Multidimensional model of racial identity: A reconceptualization of African American racial identity. *Personality & Social Psychology Review, 2*(1), 18-39.
- Sellers, R. M., & Shelton, J. (2003). The role of racial identity in perceived racial discrimination. *Journal of Personality and Social Psychology, 84*(5), 1079-1092.
- Simmons, C. M. (2013). *Reducing stereotype threat in academically at-risk African-Americans students: A self-affirmation intervention* (Order No. AAI3498889). Available from PsycINFO. (1322255048; 2013-99010-209). Retrieved from <http://search.proquest.com/docview/1322255048?accountid=14553>

- 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.
- 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.
- Steele, C.M. (1997). A threat in the air: How stereotypes shape intellectual identity and performance. *American Psychologist, 52*, 613–629.
- Steele, C.M., Spencer, S., Nisbett, R., Hummel, M., Harber, K., Schoem, D., & Carter, K. (1998). *African American college achievement: A “wise” intervention*. Unpublished manuscript, Stanford University, Stanford, CA.
- Steele, C. M. (2010). *Whistling Vivaldi and other clues to how stereotypes affect us*. New York, NY: Norton.
- Taylor, V. J., & Walton, G. M. (2011). Stereotype threat undermines academic learning. *Personality and Social Psychology Bulletin, 37*(8), 1055-1067.
- The College Board and Educational Testing Service (2009). *The Official SAT Study Guide* (2nd ed.). New York, NY: Educational Testing Service.
- Tokar, D. M., & Fischer, A. R. (1998). Psychometric analysis of the Racial Identity Attitude Scale. *Measurement and Evaluation in Counseling and Development, 31*, 138–149.
- Trochim, W. M., & Donnelly, J. P. (2006). *The research methods knowledge base* (3rd ed.). Cincinnati, OH: Atomic Dog.
- Vandiver, B. J., Cross, W. E., Jr., Fhagen-Smith, P. E., Worrell, F. C., Swim, J., & Caldwell, L. (2000). *The Cross Racial Identity Scale*. State College, PA: Author.

- Vandiver, B. J., Cross, W. E. Jr., Worrell, F. C., & Fhagen-Smith, P. E. (2002). Validating the Cross Racial Identity Scale. *Journal of Counseling Psychology, 49*, 71-85.
- Worrell, F. C., Vandiver, B. J., & Cross, W. E., Jr., (2004). *The Cross Racial Identity Scale: Technical manual – 2nd. edition*. Berkeley, CA: Author.
- Walton, G. M., & Cohen, G. L. (2003). Stereotype lift. *Journal of Experimental Social Psychology, 39*, 456–467.
- Walton, G.M., & Cohen, G.L. (2007). A question of belonging: Race, social fit, and achievement. *Journal of Personality and Social Psychology, 92*, 82–96.
- Walton, G. M., & Spencer, S. J. (2009). Latent ability: Grades and test scores systematically underestimate the intellectual ability of negatively stereotyped students. *Psychological Science, 20*(9), 1132-1139.
- Wax, A. (2009). Stereotype threat: A case of overclaim syndrome? In C. H. Sommers (Ed.), *The science on women and science* (pp. 132-169). Washington D. C.: AIE Press.
- Wright-Adams, V. (2014). *Stereotype threat and the academic task performance of African American women* (Doctoral dissertation). Retrieved from Psycinfo. (2014-99210-064)
- Yanico, B. Y., Swanson, J. L., & Tokar, D. M. (1994). A psychometric investigation of the Black Racial Identity Attitude Scale—Form B. *Journal of Vocational Behavior, 44*, 218–234.

APPENDIX A: DEMOGRAPHIC SURVEY

Directions. Please tell us about yourself by completing the following information:

1. Age: _____
2. Gender: _____ Male _____ Female _____ Transgender
3. Year in college:
 - a. Freshman
 - b. Sophomore
 - c. Junior
 - d. Senior
4. Please indicate your ACT or SAT scores:
 - a. ACT mathematics _____
 - b. ACT English _____
 - c. SAT mathematics _____
 - d. SAT Writing and Language _____
 - e. N/A
5. What is your racial or pan-ethnic identification?
 - a. Asian/Asian American
 - b. Black
 - c. Latino(a)/Hispanic
 - d. White
 - e. Other (please specify _____)

APPENDIX A: DEMOGRAPHIC SURVEY (CONTINUED)

6. Please indicate your primary ethnic background (e.g., African American, Filipino, Chinese, Taiwanese, French, Mexican American, Italian American, Haitian, Irish American, Cuban, etc.): _____
7. Please indicate the highest education level of your parent(s) (guardian(s) growing up) below:

Mother (female guardian growing up)

- a. some high school
- b. high school diploma or equivalent
- c. some college
- d. Bachelor's degree
- e. Master's degree
- f. Doctoral or professional degree (e.g., M.D., Ph.D.)
- g. other _____

Father (male guardian growing up)

- a. some high school
- b. high school diploma or equivalent
- c. some college
- d. Bachelor's degree
- e. Master's degree
- f. Doctoral or professional degree (e.g., M.D., Ph.D.)
- g. Other _____

APPENDIX A: DEMOGRAPHIC SURVEY (CONTINUED)

8. Which religion or spiritual beliefs do you identify with?
- a. Christian (e.g., Catholic, Protestant, AME)
 - b. Jewish
 - c. Hindu
 - d. Muslim
 - e. Atheist/Agnostic
 - f. Other _____
9. Please list 3 available times and days over the next two weeks that you are available for at least 30 minutes to complete Part Two of this study. This session will include the completion of a questionnaire, debriefing, and our token of appreciation (\$5 cash).
- a. _____
 - b. _____
 - c. _____

APPENDIX B: AFRICAN AMERICAN CULTURAL ENGAGEMENT

Use the numbers below to indicate how much you agree or disagree with each statement.

0	1	2	3	4
never	less than once a month	about once a month	about two or three times a month	once a week or more

1. I participate in events or activities sponsored by groups reflecting my own cultural heritage.
2. I rely on racial/cultural groups as my main support group on campus.
3. I am an active member of an African American organizational club/group.
4. I attend lectures/talks on campus about African Americans.
5. I consume information about African Americans via social media, books, articles, etc.,
6. I share information related to African Americans with others via social media, text messages, and/or conversations.

APPENDIX C: AFRICAN AMERICAN WOMEN CULTURAL ENGAGEMENT

Use the numbers below to indicate how much you agree or disagree with each statement.

0	1	2	3	4
never	less than once a month	about once a month	about two or three times a month	once a week or more

1. I am an active member of an African American **women** organizational club/group.
2. I attend lectures/talks on campus about African American **women**.
3. I consume information about African American **women** via social media, books, articles, etc.,
4. I share information related to African American **women** with others via social media, text messages, and/or conversations.

APPENDIX D: CONTROL GROUP (LOW THREAT)

"As you probably know, math skills are crucial to performance in many important subjects in college. Yet surprisingly little is known about the mental processes underlying math ability. This research is aimed at better understanding what makes some people better at math than others. But thus far, there is not a good explanation for this. The research you are participating in is aimed at better understanding these differences."

APPENDIX E: EXPERIMENTAL SCRIPT (HIGH THREAT)

"As you probably know, math skills are crucial to performance in many important subjects in college. Yet surprisingly little is known about the mental processes underlying math ability. This research is aimed at better understanding what makes some people better at math than others. As you also may know, at some top schools, White and Asian students outnumber the Black students in math majors and majors with math as a prerequisite, and there seems to be a growing gap in academic performance between these groups. A good deal of research indicates that White and Asian men consistently score higher than Blacks, especially Black women on standardized tests of math. Research also shows that White and Asian women perform better than Black women. But thus far, there is not a good explanation for this. The research you are participating in is aimed at better understanding these differences. Your test performance will be compared to other students from across the nation. One specific question is whether Whites and Asians are superior at all types of math problems or only certain types."

APPENDIX F: DEBRIEFING SCRIPT

Thank you for your participation in this experiment. While the study was a survey including a very difficult math task, I was also interested in learning about how some of you responded to comparisons to stereotypes about certain races and genders in a performance situation. Half of the participants in the study are subjected to a racial and academic stereotype about academic performance while the other half of the participants are not subjected to this stereotype. Specifically, the script that half of you were read contained language that identified race and gender as weaknesses in mathematics. Research has shown that stereotyping is often very subtle and there is continuing studies that have provided a wealth of knowledge to determine its negative effects on students. However, in spite of this, there is little research that examines cultural engagement, racial identity, experimenter race as buffers, or what psychologists refer to as moderators that reduce this negative effect. The study you participated in will increase new knowledge in this area.

One of the reasons I used deception in this study was that I was concerned if I told you I was also studying how these beliefs influence your response in performance conditions, you might on some level adjust your responses when the goal of the study is to learn about people's natural, unadjusted responses. The benefits of participating in this study is that it will increase awareness of some of the issues, such as stereotyping, that may affect students at this university or other universities where Black or African American students are in the minority. The goal of the study is to increase awareness about social and self-perception issues that could diminish academic achievement and task performance at predominately White universities.

APPENDIX F: DEBRIEFING SCRIPT (CONTINUED)

If you have any questions about this study or if you are interested in the results of the study, please email Arielle Brown at aabrown3@illinois.edu. Again, I want to thank you for your participation. It is important that you do not share the details of the study with anyone as this might compromise the results. Thank you. (*hands participant token of appreciation*)

APPENDIX G: EXPERIMENTER PROTOCOL

****Be sure to read the scripts verbatim****

PREPARATION

1. Check **Google calendar** to check condition (low threat [0] or high threat[1]) and form of compensation (cash [\$] or 1-hour EPSY/PSYC credit [credit])
2. Grab packet (in binder; see checklist)
3. Place materials in their appropriate places
 - a. Pencils are in pencil holder where participant will sit
 - b. The rest of the materials are with you and out of the sight of the participant
 - c. Hang “do not disturb” sign on door
4. Check lobby to check if participant arrived
 - a. More than 15 minutes late = mark as no-show
 - b. If no-show → email/text Arielle

GREETING PARTICIPANT

5. Greet participant in the 188 suite lobby and walk them in the room 188D (Dawn and Helen) or 188H (Arielle and Cass).

ADMINISTERING MATERIALS & MATH TASK

Please stand when reading scripts

6. Ask them to **silent their phones** (“To minimize distractions, may you please turn your phone off or put it on silent, not vibrate. Thank you.”)
7. Administer the **Math Importance Survey** (“Please complete this. Thank you.”)
8. Read the **script**—low (0) or high threat (1)

APPENDIX G: EXPERIMENTER PROTOCOL (CONTINUED)

9. Set **timer** for 25 minutes. If they finish early, go to next step.
10. While the person is testing, **record** behavior observations on table. (You could be signing extra credit form or reading after writing observations)
11. When the timer goes off say, “Stop. Put your pencil down” and collect math task.
Record math completion time on table

ADMINISTERING ADDITIONAL MATERIALS & COMPENSATION

12. Hand the participants **the test difficulty and effort questions**
13. Hand the participants the **ACT/SAT release form** with **pen!** (“This gives us permission to obtain your ACT/SAT scores from the registrar’s office”)
14. Read **debriefing** sheet
15. Sign sheet for EPSY research credit or give cash
16. Hand them Resources Sheet

STORING MATERIALS

17. Place completed math tasks in “Completed Math Tasks” folder in cabinet
18. Place completed math importance questions in “Completed Math Importance” folder in file cabinet
19. Place completed test difficulty questions in “Test Difficulty” in cabinet
20. Place observation sheet in your binder (leave binder in your room)
21. Place ACT/SAT release form in envelope and slide it under Helen’s door or hand them to her (if present). They will be collected daily and placed in a secure location.
22. Take down do not disturb sign on door

APPENDIX H: EXPERIMENTER CHECKLIST

1. Attendance sheet
2. Two pencils with erasers
3. Scripts
4. Math task packet
5. Timer
6. Math importance questions
7. Difficulty and effort questions
8. ACT waiver release form
9. Debriefing script
10. Tokens of appreciation (in drawers)
 1. \$5 cash
 2. 1-hour EPSY/PSYC research credit (be sure to sign EPSY form)
 3. Snack (on shelf)

APPENDIX I: DESCRIPTIVE STATISTICS

Table 5: Descriptive statistics for all continuous variables.

Variables	<i>N</i>	Range	<i>M</i>	<i>SD</i>	α
Age	101	18-25	19.66	1.35	
Math Task	103	0-6	3.26	1.22	
ACT Math	103	16-33	23.55	3.48	
RI Beliefs ^a	102	4-7	5.97	.80	.81
RI Behaviors ^b	102	.32-4.00	2.22	.96	.82

^a Mean scores for racial identity (RI) beliefs on the Cross Racial Identity Scale (CRIS) Internalization Multiculturalist Inclusive subscale (IMCI). ^b Mean scores for racial identity (RI) behavior on cultural engagement items. Math task represents score on math task. ACT math is ACT math scores from the registrar's office.

APPENDIX J: CORRELATION MATRIX

Table 6: Pearson product-moment correlations among continuous variables

	1	2	3	4	5	6	7
1. Math Task	–						
2. Age	-.01	–					
3. ACT math	.29**	-.02	–				
4. Math Course	.28*	-.23*	.25*	–			
5. Exper. Race ^a	-.17	.05	-.21*	.12	–		
6. RI Beliefs ^b	.18	.03	.05	-.17	-.05	–	
7. RI Behavior ^c	.19	.18	.07	-.07	-.08	.18	–

Note. Correlation is significant at the 0.05 level (2-tailed). * Correlation is significant at the 0.01 level (2-tailed). **

^a Random assignment to Black or White experimenter

^b Mean scores for racial identity beliefs on the Cross Racial Identity Scale (CRIS) Internalization Multiculturalist Inclusive subscale (IMCI). ^c Mean scores for racial identity behavior on cultural engagement items

APPENDIX K: MAIN ANALYSES WITH LISTWISE DELETION METHOD FOR MISSING ACT SCORES

All of the following analyses were conducted without the eight participants who had missing ACT math scores ($N = 95$). The missing data procedure, listwise deletion was used to compare results to EM and manual sample mean imputation. None of the results were significantly different from the findings addressed in the main text.

EFFECTS OF STEREOTYPE THREAT ON MATH PERFORMANCE

An ANCOVA was conducted to examine math task scores between high and low threat stereotype conditions. The results did not reveal a main effect by threat condition after controlling for ACT math scores ($F_{1, 94} = 1.00, p = .32$).

EXPERIMENTER RACE ON MATH PERFORMANCE

A 2 x 2 ANCOVA was conducted. No main effect for experimenter race was found ($F_{1, 94} = 1.23, p = .27$). Also, the interaction between stereotype threat condition and experimenter race was not significant ($F_{1, 94} = .43, p = .52$).

RACIAL IDENTITY ON MATH PERFORMANCE

The results did not reveal a main effect for racial identity beliefs or behaviors in the model (see Table X). No interaction terms significantly predicted math task performance. The model with all predictors (model 3) only accounted for approximately 8% of the variance in math task scores. ACT math scores were a significant predictor in every model.

**APPENDIX K: MAIN ANALYSES WITH LISTWISE DELETION METHOD
FOR MISSING ACT SCORES (CONTINUED)**

Table 7: Regression analysis on math task after listwise deletion for missing ACT scores.

<i>Variable^a</i>	B	SE B	β	<i>p</i>
Model 1				
ACT	.11	.03	.33**	.00
Stereotype threat	.25	.25	.10	.30
Model 2				
ACT	.11	.03	.33**	.00
Stereotype threat	.24	.25	.10	.35
RI Beliefs	.01	.16	.00	.97
RI Behaviors	-.05	.13	-.04	.68
Model 3				
ACT	.12	.03	.35**	.00
Stereotype threat	.23	.25	.09	.37
RI Beliefs	-.01	.25	-.01	.97
RI Behaviors	-.30	.19	-.24	.11
Ster. Threat x RI Beliefs	-.04	.34	-.02	.91
Ster. Threat x RI Behaviors	.46	.26	.27	.08

Note. P-value is significant at the 0.05 level (2-tailed). * P-value is significant at the 0.01 level (2-tailed). **

^a Covariate (ACT) and all independent variables centered on mean. Stereotype threat condition was dummy-coded (0—low threat; 1—high threat). Racial identity beliefs (RI Beliefs) were measured with CRIS IMCI subscale. Racial identity behaviors (RI Behaviors) were measured with cultural engagement items.

APPENDIX L: ADDITIONAL MEASURES

BACKGROUND

Other factors have been associated with performance with stereotype threat such as domain identification (Nadler & Komarraju, 2016; Nguyen & Ryan, 2008; Steele, 2010; Walton & Cohen, 2003), perceived difficulty, and effort (Keller, 2002). In addition to these variables, the current study also included math course enrollment as a measurement of experience in the domain. Examining additional variables' relationship to math task performance could increase understanding about individual factors that might influence one's performance in threatening conditions.

MEASURES

Math identification items. The brief, three-item survey consisted of items regarding math importance and identification for participants. It is based on the notion that domain identity is connected to one's abilities in a particular domain and self-concept (Aronson et al., 1999; Steele & Aronson, 1995). A sample item includes: "If I score poorly on math tests, I will feel badly about myself." Items are rated on a 5-point Likert-type scale ranging from 1 (strongly disagree) to 5 (strongly agree). They were adopted from Steele and Aronson's (1995, p. 807) and McFarland and colleagues' (2003, p. 190) studies.

Difficulty, effort, and performance items. I used three items from Keller's (2002) study examining perceived difficulty, effort, and performance of participants. Item scores range from 1 (I did not put much effort into it) to 7 (I put much effort into it). A sample item measuring effort: "How hard did you try? Please indicate your level of effort on this task."

APPENDIX L: ADDITIONAL MEASURES (CONTINUED)

Math course enrollment. Participants were asked to indicate their math course enrollment. Options ranged from “never taken a math course” to “currently enrolled in math.”

RESULTS

Overall, Pearson product-moment correlation was used to examine if math identification, perceived effort and difficulty, and math course enrollment were related to math task performance. Moreover, associations between the variables were assessed. Results indicated that only math course enrollment was associated with math task performance. Significant relationships between the variables were found. Independent samples *T*-tests were also conducted to assess variables between low and high threat groups and experimenter race; no significant mean differences were found. Refer to Table 9 for correlations among all additional variables.

Math identification. Pearson product-moment correlation revealed there was not a statistically significant positive relationship between math task performance and math identification ($r[101] = .11, p = .27$). In other words, math task performance was not well-predicted by how much a participant identified with math. Although math identification was not related to math task performance, it was positively associated with ACT scores ($r[101] = .22, p = .03$). Students with higher ACT math scores found math to be more important to their identity. Math identification was also correlated with effort (see next section). The alpha coefficient for the sample was .58.

APPENDIX L: ADDITIONAL MEASURES (CONTINUED)

Perceived difficulty, effort, and performance. The three items adopted from Keller (2002) were used to measure self-reported difficulty, effort, and performance after taking the math task. None of the items were significantly associated with math task performance, but there were significant associations with ACT math scores. Because these items were administered after the stereotype threat cue and math task, score differences on these items between threat conditions and experimenter race were assessed using independent samples *t* tests.

The first item that measured difficulty, “Did you have any difficulties in answering the questions?” was negatively and significantly correlated with ACT math scores ($r[101] = -.35, p < .001$). This indicated that students who scored higher on the math portion of the ACT found the math task to be less difficult than those with lower ACT math scores. Perceived difficulty was also negatively associated with performance confidence at a significant level ($r[101] = -.69, p < .001$). Participants who found the math task more difficult, indicated they performed worse than the average performance on the math task. When examining threat conditions, differences in perceived difficulty scores were not significant (Low— $M = 4.67, SD = 1.66$; High— $M = 5.00, SD = 1.51, p = .28$). Regarding experimenter race, perceived difficulty did not significantly differ between participants randomly assigned to a Black experimenter or White experimenter (Black— $M = 4.69, SD = 1.59$; White— $M = 4.98, SD = 1.58, p = .84$).

Effort (i.e., “How hard did you try?”) was significantly and positively correlated with math identification. Individuals who indicated they identified more with math, indicated they gave more effort on the math task. The independent samples *t*-test

APPENDIX L: ADDITIONAL MEASURES (CONTINUED)

indicated the means were not significant between low and high threat groups (Low— $M = 5.78$, $SD = .92$; High— $M = 6.06$, $SD = .98$, $p = 1.00$). This was also seen when grouped by experimenter race (Black— $M = 6.02$, $SD = .92$; White— $M = 5.82$, $SD = .99$, $p = .26$). Therefore, participants' effort was similar across threat groups and experimenters, White or Black.

Individuals with higher performance confidence (i.e., “How well do you think you performed in answering the questions compared to the average performance?”) had higher ACT math scores, identified more with math, and reported less difficulty with the math task. Similar to the difficulty and effort items independent samples t -test results, performance confidence was not significantly related to threat condition (Low— $M = 3.75$, $SD = 1.66$; High— $M = 3.54$, $SD = 1.54$, $p = .51$) or experimenter race (Black— $M = 3.77$, $SD = 1.52$; White— $M = 3.51$, $SD = 1.68$, $p = .33$).

Math Course enrollment. To assess math task performance's relationship math prior experience, participants were asked when their last math course was taken. Math course enrollment was positively associated with the following variables at a significant level: math task, ACT math scores, and math identification (see Table 9). Math course enrollment was negatively and significantly correlated with age, $r(101) = -.23$, $p = .02$. T -tests showed that math course enrollment did not significantly differ for low and high threat groups (Low— $M = 3.04$, $SD = 1.46$; High— $M = 3.48$, $SD = 1.44$, $p = .93$) or Black and White experimenters (Black— $M = 3.10$, $SD = 1.49$; White— $M = 3.43$, $SD = 1.42$, $p = .90$). Supporting random assignment, math course enrollment was not significantly correlated with threat condition ($r[101] = .15$, $p = .12$).

APPENDIX L: ADDITIONAL MEASURES (CONTINUED)

DISCUSSION: DID ADDITIONAL FACTORS MATTER?

Contrary to previous stereotype threat research with women and math performance (Nguyen & Ryan, 2008), math task scores were not significantly related to math identity for Black women. Meaning, women who identified more with math were not affected more or less by stereotype threat than those who did not care about math and their performance. Numerous studies have documented that stronger domain identities predicted more debilitating effects of stereotype threat (e.g., Nguyen & Ryan, 2008; Steele, 2010). This did not hold true in the current study.

Instead, math task scores were associated with previous math performance and experience, such as ACT math scores and math course enrollment. Students with higher ACT math scores and students who had taken a math course more recently, performed better on the math task. Prior research has found ACT and SAT scores to be significant covariates for math tasks, highlighting their importance as control variables and their positive relationship with task performance (Davis et al., 2006; Marx & Goff, 2005). It is intuitive that those who are more skilled in a domain will perform better on a task for that domain, regardless of threatening conditions. Controlling for performance and experience could minimize type I errors.

APPENDIX L: ADDITIONAL MEASURES (CONTINUED)

Table 8: Descriptive statistics for additional variables.

Variables	<i>N</i>	Possible Range	<i>M</i>	<i>SD</i>	α
Math Identity	103	1-5	3.73	.61	.58
Difficulties	103	1-7	4.83	1.58	
Effort	103	1-7	5.92	.96	
Performance	103	1-7	3.64	1.60	

Note. Math Identity indicates mean scores on the math identification measure. Difficulties represents the mean score X items about how difficult they found the math task. Effort indicates mean score on effort question.

Table 9: Pearson product-moment correlation matrix for additional variables

	1	2	3	4	5	6	7	8
1. Math Task	–							
2. Age	-.01	–						
3. ACT math	.29**	-.02	–					
4. Math Identity ^a	.11	-.16	.22*	–				
5. Difficulties	-.06	.19	-.35**	-.14	–			
6. Effort	.03	.10	-.05	.31**	.08	–		
7. Performance	.16	.07	.41**	.21*	-.69**	.08	–	
8. Math Course	.28**	-.23*	.25*	.32**	-.03	-.01	.07	–

Note. Correlation is significant at the 0.05 level (2-tailed).* Correlation is significant at the 0.01 level (2-tailed).**^a Mean score for math identification items. Difficulties, effort, and performance were all measured using one item.

APPENDIX M: IRB APPROVAL

UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN

Office of the Vice Chancellor for Research

Office for the Protection of Research Subjects
528 East Green Street
Suite 203
Champaign, IL 61820



August 12, 2016

Helen Neville
Educational Psychology
226 Education Bldg
1310 South Sixth Street
Champaign, IL 61820

RE: *Student Beliefs and Academic Performance in College*
IRB Protocol Number: 17032

Dear Dr. Neville:

This letter authorizes the use of human subjects in your project entitled *Student Beliefs and Academic Performance in College*. The University of Illinois at Urbana-Champaign Institutional Review Board (IRB) approved, by expedited review, the protocol as described in your IRB application. The expiration date for this protocol, IRB number 17032, is 08/08/2017. The risk designation applied to your project is *no more than minimal risk*.

Copies of the attached date-stamped consent form(s) must be used in obtaining informed consent. If there is a need to revise or alter the consent form(s), please submit the revised form(s) for IRB review, approval, and date-stamping prior to use.

Under applicable regulations, no changes to procedures involving human subjects may be made without prior IRB review and approval. The regulations also require that you promptly notify the IRB of any problems involving human subjects, including unanticipated side effects, adverse reactions, and any injuries or complications that arise during the project.

If you have any questions about the IRB process, or if you need assistance at any time, please feel free to contact me at the OPRS office, or visit our website at <https://www.oprs.research.illinois.edu>.

Sincerely,

A handwritten signature in black ink that reads "Ronald A. Banks".

Ron Banks, MS, CIP
Human Subjects Research Coordinator, Office for the Protection of Research Subjects

Attachment(s): Informed Consent Documents; Waiver of Documentation of Informed Consent

c: Arielle Brown

APPENDIX M: IRB APPROVAL (CONTINUED)

UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN

Office of the Vice Chancellor for Research
Office for the Protection of Research Subjects
805 West Pennsylvania Ave
Urbana, IL 61801



July 28, 2017

Helen Neville
Educational Psychology
226 Education Bldg
1310 South Sixth Street
Champaign, IL 61820

RE: *Student Beliefs and Performance in College*
IRB Protocol Number: 17032

Dear Dr. Neville:

This letter authorizes the use of human subjects in your continuing project entitled *Student Beliefs and Performance in College*. The University of Illinois at Urbana-Champaign Institutional Review Board (IRB) approved the protocol as described in your IRB application, by expedited continuing review. The expiration date for this protocol, IRB number 17032, is 07/26/2020. The risk designation applied to your project is *no more than minimal risk*.

Copies of the attached date-stamped consent form(s) must be used in obtaining informed consent. If there is a need to revise or alter the consent form(s), please submit the revised form(s) for IRB review, approval, and date-stamping prior to use.

Under applicable regulations, no changes to procedures involving human subjects may be made without prior IRB review and approval. The regulations also require that you promptly notify the IRB of any problems involving human subjects, including unanticipated side effects, adverse reactions, and any injuries or complications that arise during the project.

You were granted a three-year approval. If there are any changes to the protocol that result in your study becoming ineligible for the extended approval period, the RPI is responsible for immediately notifying the IRB via an amendment. The protocol will be issued a modified expiration date accordingly.

If you have any questions about the IRB process, or if you need assistance at any time, please feel free to contact me at the OPRS office, or visit our website at <https://www.oprs.research.illinois.edu>

Sincerely,

A handwritten signature in blue ink that reads "Michelle Lore".

Michelle Lore, MS
Human Subjects Research Specialist, Office for the Protection of Research Subjects

Attachment(s): 2 Consent Forms

c: Arielle Brown