# MEASUREMENT INVARIANCE IN MATH ANXIETY SCALES ACROSS RACE AND GENDER

---

A Dissertation
Submitted to
the Temple University Graduate Board

---

In Partial Fulfillment
of the Requirements for the Degree
DOCTOR OF PHILOSOPHY
SCHOOL PSYCHOLOGY

---

by
Mariah N. Davis
May 2023

Examining Committee Members:

Laura L. Pendergast, Advisory Chair, School Psychology Department
Lia E. Sandilos, School Psychology Department
W. Joel Schneider, School Psychology Department
Kelly McGinn, External Member, Human Development and Community Engagement

**ABSTRACT**

This study investigated the reliability and validity of the Abbreviated Math Anxiety Scale (AMAS; Hopko, Bare, & Hunt, 2003) and the Math Anxiety Scale – Revised (MAS-R; Bai, 2009) across race and gender as well as the extent to which race and gender predict learning math anxiety and math evaluation anxiety at both the secondary level, which includes middle and high school students, and the tertiary level, which includes college/university students. Cronbach's Alpha scores were compared across race and gender subsamples, confirmatory factor analyses were conducted for the theoretically aligned structures of the AMAS and MAS-R, multi-group confirmatory factor analyses were conducted across race and school level for the AMAS, and hierarchical regression analyses were conducted with learning math anxiety and math evaluation anxiety as predicted variables and race and gender as predicting variables.

Findings from this study suggest that the AMAS and MAS-R are sufficiently reliable across race and gender subgroups, that the theoretical factor structure of the AMAS fits the data from this study, while that of the MAS-R does not, and that the AMAS is non-invariant across White/European American and Black/African American racial groups as well as secondary level and tertiary level students. Results of the hierarchical regression analysis indicate that race and gender are not predictive of learning math anxiety or math evaluation anxiety.

Taken together, these results suggest that math anxiety may not be a significant factor in the development of differences in attrition and retention rates across race and gender in Science, Technology, Engineering, and Mathematics (STEM) programs, but that the AMAS is a valid measure to use with Black/African American students, secondary level students, and tertiary level students. Implications and future directions are discussed.

## DEDICATION

This work is dedicated to a variety of individuals who have collectively made up my strong, unwavering support system. First, to my friends and peers, Stephanie Joseph, Samantha Rushworth, Kaiyla Darmer, and Johnson Ho, with whom I shared many long nights and weekends working through various trials and tribulations throughout our educational journeys. Second, to my parents, Lisa and John Davis, who have always supported me with all of my aspirations and life goals, and who root for me every single day. Third, to my fiancé, Paul Lavelle, ever-patient and loving throughout the dissertation process, who always made sure I ate dinner at a reasonable hour, even on my busiest days. Finally, to my dissertation chair, Laura Pendergast, and my committee, Lia Sandilos, W. Joel Schneider, and Kelly McGinn, who supported my development as a researcher and scholar, provided invaluable contributions to my work, and saw me through to completion of my doctoral dissertation and program.

**TABLE OF CONTENTS**

## LIST OF TABLES

# CHAPTER 1

# INTRODUCTION

Technology has been advancing rapidly in recent years, leading to quicker societal and economic changes, and Science, Technology, Engineering, and Mathematics fields, commonly referred to as "STEM" fields, are a driving force for new innovations in our society (National Science Foundation, 2020). Due to such rapid innovation, the number of STEM careers are increasing, and some have raised concerns that the demand of qualified applicants to STEM careers outweighs the supply (Iammartino et al., 2016; Mau & Li, 2018; Xue & Larson, 2015). Retention rates in college degrees that lead to STEM careers are low in general, but they are particularly low for ethnic minority students (Major et al., 2012). Similarly, women and racial minority groups are underrepresented in STEM fields, with the exception of Asian Americans (National Science Foundation, 2019). One might attribute these low retention rates and representation in STEM programs and fields to lack of ability, but research suggests that academically talented women and minority students often choose not to pursue careers in STEM (Kerr & Kurpius, 2004; McCormick, 1998). The choice of academically talented women and minority students to not pursue a career in STEM suggests that there is a non-academic reason for the low interest and retention rates of marginalized gender and racial groups in STEM fields. Social-emotional factors, including anxiety, may be considered as possible causes of low interest and retention rates for STEM fields and programs in marginalized racial and gender groups. One type of anxiety that is prevalent among students and impacts both interest and performance in STEM-related subjects is math anxiety.

Math anxiety refers to the fear, tension, or apprehension experienced when performing mathematics, learning mathematics, or being evaluated in mathematics (Ashcraft, 2002). Math anxiety often causes a decline in performance when working math problems under timed, high-stakes conditions (Ashcraft & Moore, 2009). Additionally, individuals with math anxiety are less likely to pursue math classes, math learning programs, and math-related careers (Ashcraft & Moore, 2009; Ramirez et al., 2018). Qualitative research from a sample of women in college who come from marginalized and/or underserved racial, economic, and education backgrounds has found that stereotypes about women in math negatively influence self-perceptions of math ability, which can in turn lead to math anxiety and poor performance (Amato, 2018). Additionally, the conviction that math is a white male domain has been found to contribute to math anxiety among college students with high levels of math anxiety (Tobias, 1990).

The absence of representation of race and gender in STEM is a grave concern. First, the exposure to representation of one's marginalized identity in STEM fields is a predictor of an individual with a marginalized identity maintaining a STEM career goal from an early age (Kerr & Kurpius, 2004). Therefore, representation is important to maintain equity in career options across marginalized groups. Second, the absence of diversity in the development of STEM products creates a problem for those who are not represented at the table in which these products are being designed. As a society, the United States is heavily reliant on internet usage (File & Ryan, 2014). Reliance on internet usage is facilitated by products created by STEM fields, such as iPhones, tablets, and laptops; which are regularly used for both work and leisure in the United States

(Horrigan, 2009). Usage of these products starts at a young age, as most children in today's society in the United States receive their first cell phones and virtual accounts when they are in middle school (Lenhart, 2010). Further, marginalized groups such as Black/African American individuals use iPhones and other mobile internet devices more frequently than White/European American individuals, and rates of their usage of mobile internet devices are growing faster (Horrigan, 2009). However, BIPOC individuals are not served well by STEM creators who are not diverse, or companies that do not have representative groups of people on their teams. Just as one example, facial recognition technology often fails for racial minority groups (Grother et al., 2019; M. Wang et al., 2019). In particular, findings suggest high rates of false positives for Black/African American women, which becomes a severe social justice concern when it is used to identify individuals in law enforcement (Grother et al., 2019). Additionally, in a test of a facial recognition algorithm released by Amazon, findings included an error rate of 39% for non-White/European American individuals, who were incorrectly matched with faces of individuals who had been convicted of crimes (M. Wang et al., 2019).

Demographic variables are important to consider for the research and development of STEM products. As of now, demographic factors are largely not considered in Human-Computer Interaction (HCI) research (Yardi & Bruckman, 2012). The absence of the consideration of demographics in the development and research of STEM products may contribute to the failure of facial recognition technology for marginalized groups. Relatedly, team diversity on creative projects leads to increases in innovation, with various demographic variables (Dayan et al., 2017; Østergaard et al., 2011). Additionally, groups of diverse problem-solvers can outperform groups of individuals solely selected for problem-solving ability (Hong & Page, 2004). Furthermore, having individuals from marginalized groups on STEM development teams gives a voice to the high percentages of marginalized product-users whose criticisms about the efficacy and real-world consequences of using these products for their groups would otherwise not be addressed in the development process.

As the entrance into STEM fields is reliant on the entrance and graduation from STEM-oriented higher education programs, recruitment and retention must be considered among individuals who are applying to such programs, or adolescents, as well as tertiary level students who are currently in such programs and/or taking STEM courses. Considering the role that social-emotional factors play in retention in math programs (Ashcraft & Moore, 2009; Ramirez et al., 2018), both in higher education and secondary school, math anxiety must be targeted in secondary and tertiary level students, to prevent a decrease in entrance to, and retention in, STEM programs in college. Interventions are needed for racial minority students and girls to tackle math anxiety before entrance into college, so that their math anxiety can be addressed before they choose their college majors, as well as while they are in college, in order to retain them in STEM programs. Some research has suggested the efficacy of mindfulness cognitive therapy (LaGue et al., 2019) and Bibliotherapy (Hebert & Furner, 1997) as math anxiety interventions for secondary education students, but not much research has evaluated their effectiveness for marginalized gender and racial groups. At the post-secondary level, mindfulness, breathing strategies, counseling, expressive writing, and growth mindset interventions have all been seen to reduce math anxiety in samples of college students (Brunyé et al., 2013; Hendel & Davis, 1978; L. Iossi, 2007; Park et al., 2014). In particular, racial minority college students have benefitted from math anxiety interventions that have included self-affirmations (Samuel & Warner, 2021). It is also possible that the intervention strategies described above may be effective as a preventative measure for adolescent girls and racial minority students who suffer from math anxiety, to combat their avoidance of math-related college majors and career paths.

However, before we can conduct these interventions, we need to make sure that we are measuring math anxiety accurately across groups. When selecting students for intervention, it is important that our measures are sensitive and specific so that we include the students who need the intervention and that we do not include students who do not need the intervention. While we want to make sure that all students get a math anxiety intervention who need it, we also do not want to take up valuable instruction time for students who do not need a math anxiety intervention. It is possible that these students are in need of a different type of intervention, and our measures are meant to aid us in discerning which interventions are needed for each case. Validity of measures is particularly important for members of marginalized groups because measures of math anxiety have historically been validated on samples without sufficient representation of such groups, including Black/African American groups (J. R. Young & Young, 2015, 2016). Lack of representation is accompanied by a lack of psychometric analyses conducted on math anxiety measures across Race. The reliability and validity studies of the standardization samples of math anxiety measures are conducted on samples that lack a representative number of racial minority cases, or samples that do not include a description of the race of participants at all. The first reports of Black/African American students in a sample using the first and widely used scale of math anxiety, the MARS

(Richardson & Suinn, 1972), did not emerge until the mid-1990s – 20 years after the MARS, and many of the first popular and widely used math anxiety measures, were validated (J. R. Young & Young, 2015, 2016).

Recent research has also concluded that the measurement of math anxiety plays a role in the understanding of how the construct affects students of diverse racial groups (J. R. Young & Young, 2015, 2016). Findings include that the reliability of the Math Anxiety Rating Scale (MARS; Richardson & Suinn, 1972), a popular measure of math anxiety, has been found to be inversely related to heterogeneity in subject population (Capraro et al., 2001). An instrument with an inverse relation to heterogeneity in subject population suggests that as study populations become more diverse, the reliability of the instrument decreases. Past research, then, utilizing the MARS, may not be consistent if utilizing a heterogenous racial sample, as is the case when a variety of racial groups are put together into one "racial minority" sample, or when a small sample of one race is compared to a significantly larger sample of another, as is frequently the case with small groups of Black/African American, or other racial minority individuals is compared to a larger White/European American sample. A more precise measurement of math anxiety with racially homogenous populations of Black/African American students is necessary to ascertain the manifestation of math anxiety in Black/African American students (J. R. Young & Young, 2015, 2016). It is possible that other measures are guilty of comparing small samples of racial minority groups to large samples of White/European American groups as well, and so the validity of other popularly used measures of math anxiety should be examined with racially diverse samples. The validity of a measure across groups can be evaluated through the examination of measurement invariance across those groups (Pendergast et al., 2017; Vijver & Tanzer, 2004). Measurement invariance investigates the degree to which a measure can effectively and accurately compare scores across groups (Vijver & Tanzer, 2004). To discern validity of math anxiety measures across race and gender, the examination of measurement invariance will be utilized in this study for math anxiety measures across race and gender to investigate measurement bias, or systematic differences between groups on an instrument designed to measure a given construct that are not due to the construct itself (Benítez et al., 2019; Vijver & Tanzer, 2004).

## Summary

This study sought to investigate the measurement invariance of two popular and widely used math anxiety scales, the AMAS (Hopko et al., 2003) and the MAS-R (Bai, 2011), across Black/African American and White/European American adolescent students. Using techniques such as Cronbach's alpha and McDonald's omega, multi-group confirmatory factor analysis (MG-CFA), and multiple regression, this study will examine the reliability and validity of the AMAS and MAS-R across race and gender, with particular consideration for configural, metric, and scalar invariance across groups as well as significant differences in average levels of math anxiety across groups. The sample aims to include 400 Black/African American and 400 White/European American adolescents; approximately half of each group will be boys and half will be girls in order to effectively examine measurement invariance across gender and race. This sample will be recruited from local and nonlocal high schools, as well as social media groups targeting parents of adolescents. Results of this study will be of interest to multiple stakeholders, including the students themselves, their parents, high school teachers, and math anxiety researchers.

# CHAPTER 2

# LITERATURE REVIEW

## Representation in STEM

Women and racial minority groups are underrepresented in STEM fields (National Science Foundation, 2019). The number of STEM careers are increasing, and some have raised concerns that the demand of qualified applicants to STEM careers outweighs the supply (Iammartino et al., 2016; Mau & Li, 2018; Xue & Larson, 2015). Retention rates in college degrees that lead to STEM careers are low in general, and particularly low for ethnic minority students (Major et al., 2012). Only 40 percent of students who enter a STEM major graduate in the United States, while underrepresented ethnic minority groups graduate at a rate of 29% (Major et al., 2012). Talented women and minority students often choose not to pursue careers in STEM (Kerr & Kurpius, 2004; McCormick, 1998). It is important for individuals with marginalized identities to see themselves represented in a career through others of a similar background, as representation is a predictor of maintaining a STEM career goal from an early age (Kerr & Kurpius, 2004). Increasing career interest in STEM fields for women and minority students may help increase the supply of qualified applicants. Additionally, it is important for diverse gender and racial/ethnic groups to be represented in STEM fields to avoid the gender and racial bias that results from their exclusion in research and development of STEM-related products.

Career interest in STEM is affected by social-emotional variables. Self-efficacy in math, or mathematics self-efficacy, is important for career interest as well as later STEM career choice (Huang et al., 2019; Zeldin & Pajares, 2016). Mathematics self-efficacy refers to a perceived confidence in one's mathematics abilities (Huang et al., 2019; Toland & Usher, 2016). Higher self-efficacy in math is also associated with lower math anxiety (Akin & Kurbanoglu, 2011; Hackett, 1985). Math anxiety refers to the fear, tension, or apprehension experienced when performing mathematics, learning mathematics, or being evaluated in mathematics (Ashcraft, 2002). Many math teachers claim that math anxiety comes from a fear of failure and a sense of inadequacy in their students (Perry, 2004). As a consequence of the experience of math anxiety, many individuals begin to exhibit avoidant behavior, taking fewer math courses and avoiding math majors and career paths (Ashcraft & Moore, 2009).

## Math Anxiety

Math anxiety became a topic of particular interest in psychology and education in the 1970s (Fuson, 2007; Hembree, 1990). The construct first emerged in the 1950s with the *Dutton Scale* (Dutton, 1954),which measured a person's feelings towards mathematics, and continued with a small number of studies in the 1960s (Fuson, 2007). Researchers then began to expand the construct of test anxiety, which was well-developed at the time, to create a more specific construct that focused solely on mathematics (Hembree, 1990). Math avoidance and math anxiety were hypothesized to stem from concerns with confidence and nerves, rather than intellect (Fuson, 2007). Early research examined gender differences in math anxiety to address the lower rates of women enrolling in STEM programs/courses in high school and college (Hembree, 1990).

The first measure of math anxiety appeared in 1972 with the creation of the first scale in the popular and widely used MARS series (Richardson & Suinn, 1972). The rationale behind creating a measure of a specific anxiety type came through research suggesting a more specific knowledge of what makes someone anxious can improve treatment outcomes. Additionally, for math anxiety in particular, research at the time had suggested that many individuals being treated for anxiety rated mathematics as the singular area in which they are anxious or tense. Early definitions of math anxiety suggest that it involves tension or worry associated with the completion of math problems or manipulation of numbers in academic settings as well as in regular daily life. Early measures examined math anxiety as a homogenous construct with a unidimensional factor (Richardson & Suinn, 1972).

More recent definitions of math anxiety are multidimensional. Factors include numerical anxiety (Suinn & Edwards, 1982) learning math anxiety (Hopko et al., 2003; Plake & Parker, 1982) and mathematics test, and/or evaluation, anxiety (Hopko et al., 2003; Plake & Parker, 1982; Suinn & Edwards, 1982). Numerical anxiety refers to anxiety surrounding the act of solving math problems in school or in everyday life. Learning math anxiety refers to students' experiences in the acquisition process of math skills. Math evaluation anxiety refers to anxiety that is exclusively experienced when one is being evaluated in math skills and/or competencies. Evaluation can be in a traditional testing environment or informally during class. In contrast, math test anxiety more specifically targets anxiety surrounding the traditional testing environment. More contemporary measurement scales have included items designed to measure math learning anxiety and math evaluation anxiety in a two-factor model of math anxiety (Hopko et al., 2003; Plake & Parker, 1982).

Some researchers have suggested that, in order to capture the extent of the construct of math anxiety, a measure must include both positive and negative affect associated with math (Bai, 2009; Betz, 1978; Watson, 1988; Kazelskis, 1998). Positive affect would reflect a degree of pleasurable interaction with the environment, whereas negative affect would reflect an adverse response to the environment.

A review of math anxiety measures utilizing databases including Google Scholar and PsychInfo revealed that the most popularly cited math anxiety measures that have validity evidence for use with adolescent populations include the Math Anxiety Rating Scale for Adolescents (MARS-A; Suinn & Edwards, 1982), Abbreviated Math Anxiety Scale (AMAS; Hopko et al., 2003), Mathematics Anxiety Scale-Revised (MAS-R; Bai, 2009), and Mathematics Anxiety Rating Scale – Revised (MARS-R; Hopko, 2003). Adolescent populations are emphasized here due to the increase in math anxiety at the secondary education level (Huang et al., 2019). While the MARS-A, AMAS, and MARS-R are designed to measure math anxiety through learning math anxiety and math evaluation anxiety, the MAS-R is designed to measure math anxiety through positive affect and negative affect.

**Causes, Effects, and Experience**

Math anxiety can emerge in various forms, ranging from moderate test anxiety to extreme anxiety including physiological symptoms such as nausea, heart palpitations, and sweaty palms (Perry, 2004). Individuals with high levels of math anxiety also differ in neural responses when presented with a math task. Research findings about differences in neural responses across levels of math anxiety when presented with a math task include increases in cortisol (Sarkar et al., 2014), higher levels of activation in the amygdala

(Young et al., 2012), decreased deactivation of areas that interfere with goal-oriented task completion (Pletzer et al., 2015), and an increase in neutral responses associated with physical pain (Lyons & Beilock, 2012). As one may deduce from these differences in neutral activity, math anxiety often causes a decline in performance when working math problems under timed, high-stakes conditions (Ashcraft & Moore, 2009). This decline in performance occurs across educational and research laboratory settings, and it leads to a discrepancy between true scores and test scores for individuals who are experiencing math anxiety both on school exams and in achievement tests designed to examine differences in achievement for research purposes (Ashcraft & Moore, 2009). The differences in performance for those experiencing high levels of math anxiety is largely due to the impact that anxiety has upon working memory, which is a critical cognitive ability necessary for problem-solving and quantitative reasoning, especially during timed tasks (Ashcraft & Kirk, 2001). The physical experience of anxiety, including a physiological tenseness, shortness of breath, and dizziness, causes the body to focus more on stimuli-driven information processing vs. goal-driven information processing (Ashcraft & Moore, 2009). In other words, anxious individuals are more inclined to focus on the perceived threat of the math task, or their worries and fears associated with it, instead of solving the written problem in front of them.

### *Exacerbation in Marginalized Students*

Research has suggested that the relationship between math anxiety and math achievement may be stronger for Black/African American students, particularly in grades 9-12 (J. R. Young & Young, 2015). Qualitative research from samples of college students with high math anxiety found that fear of mathematics, the conviction that math is a white male domain, and the conviction that one is either good in mathematics or in language arts, are all major contributors to their anxiety (Tobias, 1990). In particular, qualitative research from a sample of women in college who come from marginalized and/or underserved racial, economic, and education backgrounds found that stereotypes about women in math negatively influence self-perceptions of math ability, which in turn cause math anxiety and poor performance (Amato, 2018). College students who experience such fears, convictions, and self-perceptions also tend to lack coping skills to deal with these negative thoughts and the feelings that followed or preceded them. In addition, several college students reported that a major source of math anxiety was the classroom environment. Many such students would prefer a discussion-based class in which they could share their life experiences as opposed to a traditional styled classroom in which they feel pressured to find one right answer to every problem (Amato, 2018; Tobias, 1990). Many also exhibit self-defeating behaviors such as procrastination and not seeking help from instructors when struggling (Amato, 2018).

**Differences Across Gender and Race**

If math anxiety has a direct impact on career interest and STEM self-efficacy for girls, and women are underrepresented in STEM fields, it follows that the experience of math anxiety may be either greater or different for girls. The same may be true of racial minority groups, who also experience stereotype threat (Steele, 2010). Stereotype threat refers to the threat an individual or group of individuals feels when he or she is faced with a situation in which anything he or she does that fits a stereotype related to one of his or her identities could be taken as confirmation of that stereotype. It is of particular concern for individuals with marginalized identities. Dowker and colleagues (2016) suggested stereotype threat as a reason for gender differences in math anxiety. They also cite a study (Aronson et al., 1999) examining stereotype threat in White/European American students taking a test when primed with suggestion that individuals of Asian descent are better at math than those of European descent. The experience of stereotype threat in the math classroom could also be true for students of marginalized races in the United States such as Black/African American or Hispanic American individuals.

***Gender***

There are some discrepancies in research findings surrounding the level and experience of math anxiety among gender and racial/ethnic minority groups at the primary, adolescent, and college level. Although research consistently reveals math anxiety to be higher for women than men in adult populations (Hembree, 1990), research is mixed for gender differences in primary and secondary schools. Little research has been conducted on math anxiety in primary school samples and those that do often do not look at gender differences (Hill et al., 2016; Luttenberger et al., 2018). Furthermore, those that have looked at gender differences display inconsistent findings (Hill et al., 2016; Luttenberger et al., 2018). Some studies have found no gender differences (Punaro & Reeve, 2012; Ramirez et al., 2013), while others have found higher levels of math anxiety in girls than boys (Griggs et al., 2013; Yüksel-Şahin, 2008). In particular, math test anxiety has been found to be higher in $5^{th}$ and $6^{th}$ grade girls than boys. Some studies have measured math anxiety through the lens of attitudes towards mathematics. Such studies have found more positive attitudes towards math in boys than girls (Hill et al., 2016).

However, research has suggested that math anxiety peaks during the middle and high school years, and it can impact major career and life decisions when experienced at this age (Huang et al., 2019). Thus, math anxiety during secondary school is important to examine, as it can be linked to later STEM career choices. Secondary school studies are more consistent with studies on adult populations (Hill et al., 2016; Luttenberger et al., 2018). Many have revealed higher math anxiety in girl than boys (Buratta et al., 2019; Hill et al., 2016; Primi et al., 2014). The trend of higher math anxiety in girls can be found in international studies as well (Else-Quest et al., 2010). Although still some studies have found no gender differences (Ahmed, 2018; Birgin et al., 2010; Kyttälä & Björn, 2014). Still, even when no gender differences are found, girls' performance tends to be more highly affected by their math anxiety than boys' (Kyttälä & Björn, 2014). The research results garnered from studies finding gender differences in math anxiety at the secondary level (Huang et. al., 2019; Hill et. al., 2016; Luttenberger et. al., 2018; Buratta et. al., 2019; Primi et. al., 2014; Else-Quest et. al., 2010) suggests that the ability to accurately identify math anxiety in girls at the secondary level is important for self-efficacy as well as future career development in STEM, as girls are largely at-risk for developing math anxiety symptoms at this age.

### ***Race/Ethnicity***

A meta-analysis by Hembree (1990) examining the nature, effects, and relief of math anxiety also reviewed findings of differences in math anxiety among demographic variables, including race. Hembree (1990) suggests that there is no significant difference between White/European American and Black/African American individuals in a college sample, but that Hispanic American individuals have higher levels of math anxiety (d =

0.82, $P < .01$). However, a more recent study (Fuson, 2007) examining math anxiety across age, gender, and ethnicity in a sample of university students in the United Kingdom (UK) found no significant differences between a sample of White/European Americans and a “minority” ethnic group sample, which consisted of Black/African Americans, Asian Americans, Native Americans, and an “other” ethnic group category. Researchers chose to split ethnicity into two categories, White/European American and minority, due to a limited number of participants in ethnicity groups other than White/European American. The splitting of race/ethnicity into White/European American and minority may have contributed to the discrepancies found between Fuson (2007) and Hembree (1990). Additionally, Fuson (2007) used the *Attitudes Toward Mathematics Inventory* (ATMI; Tapia & Marsh, 2000) as a measure of math anxiety. The ATMI is used as a measure of math anxiety, but the construct is measured differently than other popular measures of math anxiety, which emphasize the worry or tension associated with mathematics. In contrast, the ATMI is measured by three factors: self-confidence, enjoyment of mathematics, and value of mathematics. Interestingly, when examined further, Fuson (2007) found that, while the White/European American group scored above the mean on the self-confidence and enjoyment of mathematics factors, the racial minority group scored below the mean on these two factors, although no statistically significant differences were found. Taken together, Hembree (1990) and Fuson (2007) suggest that, although research suggests there are not statistical differences between White/European American and Black/African American adults, there appear to be non-statistically significant differences that favor White/European Americans in terms of self-confidence and enjoyment of mathematics. Additionally, there is a discrepancy between studies as to the difference in levels of math anxiety between Hispanic American adults and adults of other races.

***Discrepancies in Secondary School.*** At the secondary level, only a handful of studies have been conducted examining math anxiety across racial groups, and most of them were conducted for doctoral dissertations. Many of these studies have found no statistically significant differences among racial groups. For example, Merritt (2011) suggests that there are no racial or gender differences among a sample of $7^{th}$ graders in a Mississippi school district. Racial groups examined in her study included Black/African American and White/European American students, and the Math Anxiety Rating Scale for Adolescents (MARS-A; Suinn & Edwards, 1982) was utilized to measure math anxiety. Clark (2004) also found no statistically significant differences between an Black/African American group and a Latinx American group on levels of math anxiety in high poverty high schools. However, she emphasizes that the average levels of math anxiety were higher for Black/African American and Latinx American students than the control group, the normative sample for the MARS-A (Suinn & Edwards, 1982). The normative sample for the MARS-A is likely majority White/European American, but data was not collected on race for participants. Additionally, socioeconomic status (SES) may be an influencer in the differences found between racial groups in this study, as the schools that the Black/African American and Latinx American students were recruited from were known to have high levels of poverty. A control group more similar in SES to the population in this study is warranted to make comparisons between groups. Similarly, Carroll (2010), utilizing the same measure of math anxiety (MARS-A), found no statistically significant differences in math anxiety across ethnic, SES, or gender subgroups in a sample of 7th grade students in Georgia. However, like Fuson (2007), her sample collapsed ethnic groups into minority and nonminority subgroups, and also provided participants access to an intervention that she cites may have had an effect on the students' math anxiety scores. Several studies examining math anxiety across racial groups had a pattern of splitting up groups into minority and majority dichotomies due to poor representation of minority groups in their overall samples. Overall, Merritt (2011), Carroll (2010), and Clark (2004) suggest that there may not be statistically significant differences across racial groups in secondary school, but there are confounding variables such as SES and participation in an intervention that interfere with findings about the relationship between math anxiety and race.

Additionally, sample sizes and representation of minority racial groups in Hembree (1990), Fuson (2007), and Carroll (2010) are suspect, which may also interfere with study results. Notably, Hembree (1990), Fuson (2007) and Carroll (2010) may not have had enough power to detect type 1 error, and yield statistically significant results with such small and varied samples of racial minority groups. When sample sizes are small, a power analysis should be conducted to determine the sample size necessary to detect type 1 error when utilize mean comparison analyses such as ANOVA (Wilcox, 1995). There is no record of a power analysis being conduct for analysis procedures utilized to compare means across race in Hembree (1990), Ma (1999), Fuson (2007) or Carroll (2010), so it is unclear if enough power was present to detect type 1 error for their analyses. ANOVA was used by used by Carroll (2010) and Fuson (2007) to declare no significant differences between minority samples and White/European American samples. Hembree (1990) was a meta-analysis of 151 studies of math anxiety, of which only 3 included representative numbers of Black/African American participants, and only 2 included Hispanic American participants. Representative numbers of Black/African American students have been described by education researchers (Young & Young, 2015) as at least 10% of the overall sample, in order to match the percentage of Black/African Americans in the public-school systems in the United States, which fluctuates from between 10% and 16% depending on school and grade level (Jackson and Howard, 2014). While it is possible that having less than 10% of participants in almost all of the studies in Hembree (1990)'s analysis could have enough power to yield statistically significant results, as a power analysis does not appear to have been calculated by Hembree (1990) to ensure adequate power in his effect size comparisons, it is also possible that there was not enough power in his analysis due to low numbers of Black/African American participants.

**Longitudinal Trajectories.** A more detailed picture of experiences of math anxiety across diverse groups can be garnered through an examination of a longitudinal study with a nationally representative sample (Ahmed, 2018), looking at trajectories of math anxiety in a diverse group of students. Ahmed (2018) was based on a socio—cognitive model of math anxiety which holds that math anxiety is a result of both personal and environmental factors. Because students of differing racial and gender groups have different experiences in school, researchers felt it was important to examine math anxiety across these groups. Ahmed (2018) found that trajectories of math anxiety in 7$^{th}$ graders fell into four separate groups: Increasing, Consistently High, Decreasing, and Consistently Low. The latter two groups were found to be predictive of later STEM career choice, suggesting that a low or decreasing math anxiety trajectory in the 7$^{th}$ grade is associated with a higher probability of a future career in STEM. Black/African American and Hispanic American individuals were more likely than White/European American individuals to fall into the Increasing and Consistently High math anxiety trajectories versus the Consistently Low and Decreasing trajectories, suggesting lower probability of entering STEM careers later in life. Similar differences were found between SES groups, as measured by parental education level, with lower parental education level as more predictive of Increasing and Consistently High math anxiety trajectories. Math anxiety in this study was measured by a two-item scale that included "Doing math makes me nervous or upset". Reliability estimates ranged from .63-.76. Differences in this study's results from those mentioned above examining race differences in math anxiety may be due to the use of this measure, which was unique to this study. Overall, this study suggests that, over time, math anxiety is more likely to increase or stay consistently high in racial minority students, while it tends to decrease or remain low for White/European American students. Notably, this study used methods that did not require the calculation of statistical significance of mean differences across groups, and instead used odds ratios to determine likelihood of groups falling into trajectory groups. As Ahmed (2018) had a nationally representative sample, it is likely that this study did not suffer from the same power restraints as Fuson (2007), Hembree (1990), and Carroll (2010).

**Cultural Experiences.** Racial differences in math anxiety may be linked to ethnicity and cultural experiences, especially for children of immigrant parents. Children from high-achieving Asian countries tend to demonstrate higher math anxiety, while children from high-achieving Western European countries tend to demonstrate low math anxiety (Dowker et al., 2016; Lee, 2009). This cross-cultural difference may manifest in differences between Asian and Western European immigrants from such countries as well as differences between Asian Americans and White/European Americans who do not have a background in the high-achieving Asian and Western European countries in question. Similar differences have also been found in adolescents' perception of their parents' math anxiety and attitudes towards mathematics across White/European American and Mainland Chinese populations, leading to more anxiety in children from Mainland Chinese populations (He, 2007). This difference in attitudes and perceptions of math anxiety across cultures may warrant more examination between racial groups, taking care to consider immigrant and acculturation status, as those with Asian cultural backgrounds may have more anxiety than those with European cultural backgrounds.

**Language.** Relatedly, language differences in cross-cultural comparisons may be a factor, and have been examined to some extent (Iossi, 2009). Iossi (2009) examined bilingual and monolingual college students in a culturally diverse urban commuter college. Findings include that bilingual and monolingual college students report similar levels of math anxiety, and a factor analysis of the Abbreviated Math Anxiety Scale (AMAS; Hopko et al., 2003) has revealed a similar two-factor solution for both bilingual and monolingual students (Iossi, 2009). Iossi (2009) included students with bilingual backgrounds in Spanish, Creole, Portuguese, and other languages. Evidence from Iossi (2009) suggests that levels of math anxiety are similar for bilingual students.

## Summary of Findings

Math anxiety is a multidimensional construct (Bai, 2011; Hopko et al., 2003; Pletzer et al., 2016; Suinn & Edwards, 1982) that refers to physical, cognitive, and affective symptoms that emerge when one is presented with a math task (Ashcraft, 2002; Ashcraft & Kirk, 2001; Ashcraft & Moore, 2009; Perry, 2004; Ramirez et al., 2018). The physical, cognitive, and affective symptoms of math anxiety affect performance in mathematics (Ashcraft & Moore, 2009; Hembree, 1990; Ramirez et al., 2013), and their effects are particularly exacerbated in marginalized racial (J. R. Young & Young, 2015) and gender (Hembree, 1990; Hill et al., 2016) groups. While findings generally suggest that women and girls are more affected by math anxiety than men and boys (Hill et al., 2016), discrepancies emerge when researchers look at race/ethnicity differences. Particularly, discrepancies in the levels of math anxiety experienced across racial groups are found in studies that examine secondary school-aged students (Ahmed, 2018; Carroll, 2010; Clark, 2004; Merritt, 2011). Some studies have also looked at math anxiety across cultures (Dowker et al., 2016; He, 2007; L. H. Iossi, 2009; Lee, 2009) and have identified differences in the experience of math anxiety across racial groups. Additionally, longitudinal research (Ahmed, 2018) suggests that math anxiety is more likely to be consistently high and increase for students who identify as a member of a minority racial group. Curiously, while Merritt (2011), Clark (2004), Carroll (2010), and Fuson (2007) do not find statistically significant differences across the levels of math anxiety across racial groups, Carroll (2010), Clark (2004), and Fuson (2007) find differences that are not statistically significant in which the majority racial group experiences less math anxiety than the minority group(s) sample. Due to low numbers of non-White/European American participants, Carroll (2010) and Fuson (2007) chose to integrate all non-White/European American students into one group and call that group the racial minority group. This integration of minority groups into one sample limits findings about differences both within and between minority racial groups and has implications for the accuracy of conclusions about the experience and level of math anxiety in individuals who identify as a member of a minority racial group. In particular, as previously mentioned, such studies may suffer from a lack of power necessary to obtain statistically significant results for members of minority racial groups. As a lack of power can mask type 1 error in ANOVA analyses (Wilcox, 1995), it is possible that the significance of mean differences across racial groups in the samples examined from Carroll (2010) and Fuson (2007) have been masked by their low sample sizes. Clark (2004) was only able to run an ANOVA across Black/African American and Hispanic American groups, as she did not have sufficient numbers of White/European American students in her sample to include a White/European American group in her analysis. However, she compared mean scores of her sample to the norms of the MARS-A, which was likely majority White/European American, and found that her sample had much lower means on the MARS-A than the norm sample. It would appear, then, that sample size is a limitation of several studies that examined math anxiety across racial groups, including Clark (2004), Carroll (2010), Fuson (2007), and Hembree (1990). As such, future research needs to be conducted examining math anxiety across racial groups with adequate numbers of each group included to run a statistical analysis with enough power to detect type 1 error.

**Measurement of Math Anxiety**

When examining the findings that are discussed in the above paragraphs, it is important to consider the methods of measurement that were used, as well as the makeup and distribution of samples for data analysis procedures. Measures of math anxiety have historically been validated on samples without sufficient representation of Black/African

American groups, as well as other racial minority groups (J. R. Young & Young, 2015, 2016). Lack of representation is accompanied by a lack of psychometric analyses conducted on math anxiety measures across race as the standardization samples lack a representative number of racial minority cases and/or do not even mention the race or ethnicity of participants. Further, discrepancies in research findings across studies can be due to a lack of reliability and/or validity present in a given measure. There is reason to suspect that measurement played a role in the discrepancies in findings of studies that examined math anxiety differences across racial groups. Namely, that some studies have found differences in levels of math anxiety across racial groups while others have not. A meta-analysis examining the relationship between math anxiety and achievement in Black/African American students grades K-12 examined measurement issues, concluding that the measurement of math anxiety plays a role in the understanding of how the construct affects Black/African American students, and likely those of other racial minority groups (J. R. Young & Young, 2015, 2016). For example, the reliability of the Math Anxiety Rating Scale (MARS; Richardson & Suinn, 1972), a popular measure of math anxiety, has been found to be inversely related to heterogeneity in subject population (Capraro et al., 2001), which suggests that as study populations become more diverse, the reliability of the MARS instrument decreases. Past research, then, utilizing the MARS, may not be consistent if utilizing a heterogenous racial sample, as is the case when a variety of racial groups are put together into one “racial minority” sample, or when a small sample of one race is compared to a significantly larger sample of another, as is frequently the case with small groups of Black/African American, or other racial minority individuals, is compared to a larger White/European American sample. As the

MARS-A is the adolescent version of the MARS measure, reliability concerns may also be present for the studies mentioned above in which the MARS-A was the primary measure of math anxiety used. Researchers conclude that a more precise measurement of math anxiety with racially homogenous populations of Black/African American students is necessary to ascertain the manifestation of math anxiety in Black/African American students (J. R. Young & Young, 2015, 2016). It is possible that other measures contain reliability concerns as well, and so the reliability and validity of other popularly used measures of math anxiety should be examined with representative numbers of racial minority groups. Young (2015, 2016) chose to focus on Black/African American students, as opposed to other racial minority groups, due to the long-standing achievement gap between Black/African American and White/European American students (National Center for Education Statistics, 2009), previous research that suggests that Black/African American students tend to frequently underperform on or avoid mathematics tasks (Business-Higher Education Forum, 2011; Hagedorn & DuBray, 2010), and the link between the development of racial identity and mathematics identity for Black/African American students (Martin, 2007). The present study also chose to focus on Black/African American students, in recognition of the unique challenges that Black/African American students face when presented with a math task, due to the perpetuation of racism and negative stereotypes about the intelligence of members of this group in the United States for hundreds of years.

## Measurement Bias and Invariance

Lack of validity in a given measure can also be due to measurement bias, or lack of measurement invariance across groups. As the studies discussed above were looking to discern differences across racial and/or gender groups, and there is a historical lack of representation of minority racial groups in validation samples of math anxiety scales, measurement bias and invariance in math anxiety scales should be examined further. Measurement bias refers to systematic differences between groups on an instrument designed to measure a given construct that are not due to the construct itself. The meaning of these differences is inconsistent within and across groups (Benítez et al., 2019). Measurement bias leads to score differences between groups that are not representative of true differences in the measured construct. In contrast, measurement equivalence, or measurement invariance, can be reached in order to compare assessment scores across groups (Guilera et al., 2013; Vijver & Tanzer, 2004).

Measurement bias includes construct bias, method bias, and item bias. The assessment of measurement bias is important in cross-group comparisons: construct validity of the measure in question depends upon invariance in responses across groups (Vijver & Tanzer, 2004). Construct bias refers to the phenomenon in which a construct is understood differently by subjects depending on which group they reside in. Construct bias can occur when there is a cultural difference between groups related to the construct. Definitions of the construct are then inequivalent. Additionally, it may be the case that different behaviors are associated with the construct. Construct bias can lead to poor sampling of relevant behavior in a given group that is associated with the target construct. Similarly, it can lead to incomplete coverage of all relevant aspects of the construct for a given group. Construct bias can be evaluated with multi-group confirmatory factor analysis (MG-CFA) (Benítez et al., 2017, 2019). MG-CFA tests configural invariance and examines similarities of loadings and factors across groups.

Method bias occurs when the methods of a study create perceived differences across groups that are not related to the target construct. Method bias encompasses sample, instrument, and administration bias. Sample bias occurs when samples differ in areas that are not related to the target variable. For example, in an assessment of cognitive intelligence, there may be intergroup differences in attention control that lead to group differences in overall scores that are not due to intelligence. If assessors did not catch this difference in attentional control, they might erroneously attribute differences in scores to group differences in intelligence, which could further lead to interventions that are not appropriate for an entire group of individuals. Additionally, a misstep in attributions of score differences could lead to stereotypes about a group that become pervasive and harmful. Instrument bias refers to differences in instrument characteristics observed across groups. For example, some groups may have more familiarity with an instrument, and receive differing scores from others in groups who are not as familiar with the instrument. More familiarity with an instrument of some groups than others can lead to the occurrence of differing response styles across groups that are not due to the target construct. Administration bias refers to bias associated with the method of administration of the instrument. Administration bias may include the physical or social circumstances of the administration, the ambiguity of instructions given by the administrator, and differences in expertise of administrators across testing sessions. Method bias can be evaluated through analysis of response styles (Benítez et al., 2019). Analysis of response styles involves examining the tendency to overuse one or more answer options in a systematic way that is unrelated to the target construct. Analysis of response styles also includes the comparison of mean differences of confounding variables such as social desirability, extremity response style, and midpoint response style between groups in responses (Benítez et al., 2017).

Item Bias or Differential Item Functioning (DIF) refers to distortions at the item level that lead to differences across groups that are not due to the target construct. Item bias occurs when items have a different psychological meaning across cultures. Ambiguity of items can also contribute to item bias. If an item has different psychological meanings across groups being examined in a study, that item is said to be a biased item. To identify item bias, many statistical procedures are used. From a statistical point of view, an item is biased if individuals from different groups with the same overall construct score do not have the same expected score on the item (Vijver & Tanzer, 2004). Differential item functioning (DIF) analysis is a procedure that analyzes items to identify item bias (Benítez et al., 2019). Equal mean scores across groups also suggest a lack of item bias, or item equivalence. MG-CFA can also be used to examine item bias through an examination of intercepts and thresholds between responses on items across groups (Pendergast et al., 2017). Another example of a popular statistical technique designed to examine item bias is the Mantel-Haenszel statistic (Guilera et al., 2013; Vijver & Tanzer, 2004).The Mantel-Haenszel statistic is used for dichotomously scored items. There is evidence to suggest that examining each of the three biases together is more helpful in the pursuit of measurement validity than any one alone (Benítez et al., 2017).

When there is a lack of measurement bias, a measure is said to possess measurement equivalence. Measurement Equivalence, or measurement invariance, suggests a lack of bias in a measuring tool across groups or time (Bauer, 2005; Guilera et al., 2013). It is also referred to as the level of measurement in which it is appropriate and valid to compare scores across groups (Vijver & Tanzer, 2004). Measurement equivalence has four levels, three of which are consistent and opposite to the three levels of measurement bias. The fourth level concerns invariance in residuals, which is not included in this review as research suggests that residual invariance may not have meaningful implications for measurement bias across groups (Brown, 2006; Chen, 2008).

The first level is structural equivalence, also referred to as construct and functional equivalence. Structural equivalence refers to a scenario in which the same construct is being measured across groups, but its measures, or the tools that are used to measure the construct, might not be equal across groups. In contrast, construct inequivalence is when an instrument measures wholly, or partially, different constructs across groups. Additionally, inequivalence may occur when the construct is associated with different behaviors or characteristics across groups (Vijver & Tanzer, 2004). MG-CFA can also be used to test for construct inequivalence through an examination of whether or not the theoretical model of the construct is a good fit for the data. In other words, if the items load on the factors that the theoretical model suggests (Pendergast et al., 2017).

The following level is known as measurement unit equivalence, or metric invariance. Metric invariance suggests that the scores of a given scale have the same unit of measurement across groups, but the origin values of the scores are different (Vijver & Tanzer, 2004). The difference in origin values leads to a shift in scale of one measure when compared to another. A commonly used example of measurement unit equivalence is the difference between Kelvin and Celsius in temperature measurement. However, to make comparisons across groups, the scale difference needs to be known, and this knowledge is uncommon in most scales that exhibit measurement unit equivalence, with Kelvin and Celsius being an exception. What can be compared, however, when one does not know the scale difference, is within group differences. Such differences can be compared across groups as well. For example, whether or not the gender differences in math anxiety in Black/African American students are similar to that of White/European American students. A test of metric invariance using MG-CFA can be conducted through an examination of the magnitude of the relationships between items and the factors they load on (Pendergast et al., 2017). Relationships with similar magnitudes across groups suggests metric invariance.

The third level of measurement invariance is full scale, or scalar, equivalence. Scalar equivalence refers to a situation in which the scores on a given measure have the same measurement unit *and* origin across groups. Scalar equivalence is the highest and suggests a lack of bias in measurement. Non-invariance at the scalar level can be influenced by social desirability, stereotypes, and use of different reference groups that change responses across groups (Pendergast et al., 2017). These influencers are particularly relevant for math anxiety measurement as gender and racial stereotypes for math performance exist, and they can lead to stereotype threat (Luttenberger et al., 2018; Steele, 2010). In addition, students of differing racial identities may have different reference groups for math anxiety which can influence their responses on measure items. Similar to structural and metric invariance, scalar invariance can be examined utilizing MG-CFA. In the examination of scalar invariance via MG-CFA, both intercepts and thresholds between items must be equivalent to declare scalar equivalence.

## Popular Measures of Math Anxiety

Discrepancies found in math anxiety research among racial groups may be due to the method of measurement used by these studies. Several studies cited above that looked at race differences in math anxiety used the MARS-A for their samples (Merritt, 2012; Carroll, 2011; Clark, 2005). The MARS series is well-known and widely used in the measurement of math anxiety, and each of these studies specifically targeted adolescents for their samples. Beyond these two given reasons for using the measure, the authors of these dissertations do not justify their use of the MARS-A. Clark (2004) cites common uses for MARS-A include screening students for placement in intervention courses designed to reduce math anxiety, counsel students, evaluate effectiveness of math programs, inform curriculum decision, and for research. If schools still use the MARS-A for screening students for intervention, curriculum decisions, and program evaluation, it is important to evaluate psychometric properties with different racial and gender groups. Reliability and validity studies were conducted on the MARS-A using the standardization sample (Suinn & Edwards, 1982). A Cronbach's Alpha score of .96 was obtained. For construct validity, correlation with math grades from GPA and standardized test scores were obtained as well as a factor analysis revealing two factors: Numerical Anxiety and Math Test Anxiety.

There was no data collected on the racial identity of the individuals in the normative sample used to validate the MARS-A (Suinn, 1982). Lack of data on racial identity is likely due to the time in which the measure was published, as the importance of race data in normative samples was not widely known at that time. Today, we know that it is important to examine race as a variable in normative samples, as to not do so may lead to measurement bias. For example, different racial groups may respond differently to items in a given measure, or express a different overall understanding of a construct. Consequently, reliability or validity studies were not conducted across different racial groups for the MARS-A (Suinn, 1982). Additionally, as previous research (Young & Young, 2015, 2016) has identified measurement concerns with the MARS-A for Black/African American students, accurate levels of math anxiety may not be measurable using the MARS-A for Black/African American students specifically.

However, other measures of math anxiety are available for use with adolescent populations and have become more popular in the literature today. Such measures are much more feasible for use in schools as they are shorter in length (Bai, 2011) and have more validity evidence gathered than the MARS-A (Hopko, 2003). For example, the Abbreviated Math Anxiety Scale (AMAS; Hopko, Bare, & Hunt, 2003) is an abbreviated version of the MARS-R (Plake & Parker, 1982), which is a revision of the MARS-A. It is made up of 9 items and has two factors: learning math anxiety and math evaluation anxiety. It has an internal consistency of .9 and Test-Retest reliability of .85. There is also a gender effect in that the mean for girls ($M=21.9$, $SD = 6.9$) is higher than that of boys ($M=19.5$, $SD=6.9$). In terms of Convergent/Divergent validity, the AMAS is Convergent with the MARS-R and divergent from math grades and number of math courses taken as well as from other measures of anxiety, including acute, chronic, and performance. There is also no difference in mean scores for ethnicity on the AMAS. However, the conclusion of mean score equivalence across ethnicity was based on a White/European American vs. Minority grouping, so differences in mean scores for differing groups of racial minority students is yet unknown. Additionally, a factor analysis on a racially homogenous sample of a racial minority group has not been conducted. The original validation sample was conducted on a population of undergraduate and graduate students. However, an Italian version of the AMAS has since been validated for use with high school students in Tuscany, Italy (Primi et al., 2014). The Italian translation of the AMAS has since been used in multiple studies with Italian samples (Buratta et al., 2019; Hill et al., 2016; Z. Wang et al., 2020). A forward translation method was used to translate the scale into Italian, which included the creation and combination of two primary Italian versions of the measure by two native Italian speakers, and a focus group of five individuals who read and revised the primary Italian version several times before obtaining a final form (Primi et al., 2014). Utilizing the final form of the Italian version of the AMAS, a confirmatory factor analysis yielded the expected two factors and the Cronbach's alphas fell within an acceptable range at .86 (CI .83-.88) for the learning math anxiety subscale, and .80 (CI .83-.89) for the evaluation math anxiety subscale. Convergent validity was also found between the scale and test anxiety, as measured with the Test Anxiety Inventory (TAI; Spielberger et al., 1978). The AMAS has also been used with adolescents in UK (Devine et al., 2012) and China (Xie et al., 2019) through similar translation methods. That being said, while reliability and validity of such translated versions of the measure have been obtained, best practices in cross-cultural translation include five stages: forward translation, synthesis, back translation, expert committee review, and pretesting (Beaton et al., 2000). It will be important to take into account that only one stage of cross-cultural translation was utilized for the Italian version of the AMAS when evaluating the reliability and validity of the AMAS with the high school population sampled in this study, to ensure that the Standard American English version of the measure is also valid for use with high school populations.

The Mathematics Anxiety Scale-Revised (MAS-R; Bai, 2009) is also a good candidate for use with high school students. The MAS-R is 14 items long and includes two factors: negative affect and positive affect. It has passable psychometric properties, including an internal consistency score of .85, test-retest reliability of .71, and inter-factor correlation of .26. It also includes positive discrimination power and convergent divergent validity with students' math achievement. It was validated on a sample of $7^{th}$ and $8^{th}$ graders with a racial breakdown of 50% White/European American non-Hispanic, 20% Black/African American, 20% Hispanic American, and 5% Asian American. The factor analysis, however, was not separated by ethnicity. In other words, it was a heterogenous examination of scale validity.

The Mathematics Anxiety Rating Scale – Revised (MARS-R; Plake & Parker, 1982; Hopko, 2003) is also a good candidate for measuring math anxiety in an adolescent population. The MARS-R scale referred to in this study is Hopko's revision of the original 24-item MARS-R by Plake and Parker (1982). Hopko's MAS-R is 12 items long and contains two factors: Learning Math Anxiety and Math Evaluation Anxiety. Hopko (2003) conducted separate confirmatory factor analysis for males and females in his validation study and found similar goodness of fit indices found. However, an examination of the factor structure of these measures has also not been conducted across racial groups in an adolescent population. It is important to investigate whether popular measures used for math anxiety are valid for diverse racial/ethnic populations. Before we can study differences in math anxiety across racial/ethnic groups, we have to make sure that the measures we are using are valid for those populations.

As previously stated, a review of math anxiety measures utilizing databases including Google Scholar and PsychInfo revealed that the most popularly cited math anxiety measures that have validity evidence for use with adolescent populations include the MARS-A, AMAS, MAS-R, and MARS-R. Again, adolescent populations are emphasized here due to the increase in math anxiety at the secondary education level (Huang et al., 2019). It would be redundant to test all four, as several are part of the same series and so have near identical items. The AMAS is the shortest of the MARS series that has most validity evidence, and MAS-R is a more recent iteration of the Fennema-Sherman math anxiety scale, and measures math anxiety through positive and negative affect, as opposed to learning math anxiety and math evaluation anxiety. As such, the AMAS and the MAS-R will be evaluated in this study.

## Key Points

Talented women and racial minority students are choosing not to pursue careers in STEM (Kerr & Kurpius, 2004; McCormick, 1998), despite a growing demand for applicants (Mau & Li, 2018). Math anxiety is associated with avoidance of mathematics courses as well as STEM-related programs in college (Ashcraft & Moore, 2009). Math anxiety has also been found to deter enrollment in higher level mathematics courses that are prerequisites to STEM programs (Huang et al., 2019). Math anxiety peaks in the middle and high school years, when math courses become more challenging (Huang et al., 2019). The identification of middle and high school girls and racial minority students with high levels of math anxiety is necessary for combating math anxiety. Measures of math anxiety have historically been validated on samples without sufficient representation of racial minority groups (J. R. Young & Young, 2015, 2016). This lack of representation is accompanied by a lack of psychometric analyses conducted on math anxiety measures across race. The reliability and validity studies that standardize these measures are often conducted on a sample that lacks a representative number of racial minority cases, or does not even mention the race of participants at all. A lack of representation of minority groups can be found in the standardization samples of the MARS-A (Suinn & Edwards, 1982), the AMAS (Hopko et al., 2003), the MAS-R (Bai, 2011), and the MARS-R (Plake & Parker, 1982). There are also discrepancies in previous research seeking to examine differences in math anxiety levels across race (Ahmed, 2018; Carroll, 2010; Clark, 2004; Fuson, 2007; Hembree, 1990; Merritt, 2011). Some find evidence for significant differences (Ahmed, 2018; Hembree, 1990) while others do not (Carroll, 2010; Clark, 2004; Fuson, 2007; Merritt, 2011). Previous research (Carroll, 2010; Fuson, 2007; Hembree, 1990) also has much lower numbers of students who identify as members of racial minority groups than as White/European American. Low sample sizes may interfere with the level of power necessary to find statistically significant differences in popular mean comparison analyses, such as ANOVA which can interfere with a study's ability to obtain accurate conclusions of mean differences (Wilcox, 1995).

Additionally, several of these studies use, or analyze studies that have used, the same measure of math anxiety, the MARS-A (Carroll, 2010; Clark, 2004; Hembree, 1990; Merritt, 2011). The studies that use a different measure hold the same definition of math anxiety (Ahmed, 2018; Fuson, 2007) suggesting that findings should lean towards similar conclusions. These discrepancies and lack of representation call into question the reliability and validity of measures of math anxiety for racial minority groups.

The most widely used and efficient math anxiety measures that are valid for use with adolescents include the AMAS and the MAS-R. In order to examine reliability and validity of these measures for different groups, measurement invariance across racial groups must be investigated. Measurement invariance refers to the level of measurement in which it is appropriate and valid to compare scores across groups (Bauer, 2005; Guilera et al., 2013; Vijver & Tanzer, 2004). Measurement invariance includes configural, metric, and scalar/threshold. Configural invariance refers to the instance in which the construct being measured is the same across groups (Vijver & Tanzer, 2004). In this study, a lack of configural invariance may be the case if different racial groups have different perceptions of what math anxiety is. If different racial groups have different perceptions of what math anxiety is, the measures being used to test math anxiety may not be wholly examining the construct as it is viewed across groups. Currently, it cannot be said with certainty that members of racial minority groups view math anxiety in the same way as members of the racial majority group in these studies, as there is a lack of equal representation of racial minority groups in these studies. Clark (2004) has a representative group of Black/African American and Hispanic American individuals in her study, but they are compared to a group of White/European Americans in a different SES group, so this is still not equal representation of groups. However, there is also not evidence to suggest that the math anxiety measures are invalid across groups. Each study included racial minority groups in some capacity and did not report findings of outliers for individuals of racial minority groups in reliability or validity tests of the measures used. It is likely that the construct is measured in a valid way across groups, with similar factor structures for each group.

Once configural invariance is confirmed, metric invariance can be examined. Metric invariance refers to the instance in which scores have the same unit of measurement across groups even if the origins are different (Vijver & Tanzer, 2004). The construct must be measured equally across groups for this to be the case. Item bias can influence metric invariance as items that function differently across groups, or DIF across groups, can lead to biases across groups, as some groups respond differently to items for reasons other than differences in the target construct (Vijver & Tanzer, 2004). Some items on these measures of math anxiety have the potential for bias due to stereotype threat experienced in the classroom by racial minorities and women. For example, "watching a teacher work an algebraic equation on the board" might load more on MEA and less on LMA in racially minority students than White/European American students because the experience of anxiety in this context might evoke similar feelings to taking a test for minorities as they both occur within the classroom, which the minority student might feel more anxious in.

If the measure has metric invariance, scalar or threshold invariance can be examined. Scalar/threshold invariance refers to the instance in which the scores on a measure have the same measurement unit and origin across groups (Vijver & Tanzer, 2004). Similar to the concerns cited above with metric invariance, stereotypes and reference group differences can influence scalar/threshold invariance. Social desirability may also be an influencer at the scalar level of invariance as stereotype threat in math/education has been observed in racial minorities and women (Luttenberger et al., 2018; Steele, 2010). Social pressures surrounding the need to present as confident, in the midst of stereotypes concerning the ability of racial minorities in math/education, may contribute to differing response styles across groups, leading to differing intercepts, or score origins, across groups. Additionally, racial minorities may have a different reference group for math anxiety than White/European American students as they are often segregated into different groups in school. Multi-group confirmatory factor analysis (MG-CFA) can be used to test each of these levels of invariance (Pendergast et al., 2017). For configural invariance, factor structure across groups must be examined. For metric invariance, the magnitude of factor loadings between items and factor across groups must be examined. For scalar/threshold invariance, the intercepts must be equivalent across groups and the probability for shifting across response options must be equivalent across groups (if the latent factor scores are comparable).

## Research Questions & Hypotheses

1) How do scores on the AMAS and MAS-R compare in reliability between samples of individuals of a single racial or gender group, including those of Black/African American and White/European American, and boys and girls, and reliability of the total sample?

H1: The AMAS and MAS-R are more reliable in race-specific samples than in the total sample.

H2: There is no significant difference in reliability of the AMAS and MAS-R for gender-specific samples as compared to the total sample.

2) Do the AMAS and MAS-R display configural invariance across racial and gender groups?

H3: The AMAS and MAS-R are invariant at the configural level across Black/African American, Asian American, Hispanic American, and White/European American groups.

H4: The AMAS and MAS-R are invariant at the configural level across male and female respondents within the same racial groups.

3) Do the AMAS and MAS-R display metric invariance across racial and gender groups?

H5: The AMAS and MAS-R are non-invariant at the metric level across groups of Black/African American and White/European American individuals due to the presence of items that do not equally measure math anxiety across groups.

H6: The AMAS and MAS-R are non-invariant at the metric level across male and female respondents within the same racial groups due to the presence of items that do not equally measure math anxiety across groups.

4) Do the AMAS and MAS-R display scalar invariance across racial and gender groups?

H7: The AMAS and MAS-R are non-invariant at the scalar level across groups of Black/African American and White/European American individuals due to stereotypes and reference groups of racial minorities that lead to differences in response styles across groups that are not due to the construct of math anxiety.

H8: The AMAS and MAS-R are non-invariant at the scalar level across male and female respondents within the same racial groups due to stereotypes and reference groups of women that lead to differences in response styles across groups that are not due to the construct of math anxiety.

5) Is there a significant and meaningful difference in math anxiety per group as represented in each measure?

H9: Math anxiety will be higher for racial minorities and girls.

# CHAPTER 3

# METHODOLOGY

## Participants

The study questionnaire yielded 3076 total responses. Qualtrics bot detection software detected that 695 of these responses were likely bots, and thus were removed. Responses with duplicate IP addresses were then removed from the sample (N = 344). The sample was then restricted to responses that came from a longitude and latitude within the United States or within 1000 kilometers from the United States, considering Qualtrics' accuracy radius of 1000 kilometers. 522 responses were removed based on location. Cases that included responses that were not in Latin script were also removed (N = 23). Responses were then split into subsamples based on education level (Secondary or College). The rationale for adding college students to the sample is described in the procedures section below. 175 responses were removed because their age did not match with the education level that they indicated. For the secondary school sample, the age range accepted was 13-18. For the College sample, the age range accepted was 16-29. Responses were further split into subsamples based on racial identity. 843 responses were removed because of inconsistencies in their reporting of racial identity. For example, many participants reported Black/African American as their race under the "Race" item but reported an ethnicity of "White" or a variation of the White/European American race or vice versa. Notably, the questionnaire had a space to indicate mixed Race or identification under multiple racial identities. After these removals, a total of 819 responses were retained. However, 65 responses were then removed because respondents indicated that they were in middle school and reported ages between 15 and 17.

The total sample after these removals was 755. The secondary school sample contained 329 total responses (39 Middle School/Junior High, 289 High School) and the college sample contained 426 total responses. One respondent indicated they were at a K-6$^{th}$ grade elementary school, and was retained in the secondary school subsample as they indicated they were 13 years old.

In terms of racial identity, 208 respondents identified as Black/African American, 450 identified as White/European American, 53 identified as Asian American, 36 identified as Native American, and 8 reported “Other” for the racial identity item. Two respondents who reported “Other” indicated that they were mixed race (Hispanic/White and Hispanic/Afro Boricua).

In terms of gender identity, 449 respondents identified as male, 294 identified as female, and 10 respondents chose the “Other” option for gender identity.

547 respondents, or 72.45% of the sample, reported that they were eligible for free or reduced priced lunch, which indicates low socioeconomic status. Parental education level tended to be similar for mothers and fathers of participants, with the majority of respondents (34-36%) indicating that their parent had earned a bachelor’s degree, followed by ~22% indicating that their parent had attended some college but did not earn a degree. Less than 20% of respondents indicated that the highest level of education that their parent had obtained was some high school, a high school degree, or a graduate degree.

Demographic characteristics of the sample are reported in Table 1.

**Table 1.** Demographic Characteristics of Sample

| Demographic | Sample Size | Percentage |
| --- | --- | --- |
| Race | | |
| Black/African American | 208 | 27.55% |
| White/European American | 450 | 59.60% |
| Asian American | 53 | 7.02% |
| Native American | 36 | 4.77% |
| Other Race | 8 | 1.06% |
| Ethnicity | | |
| Hispanic/Latinx | 318 | 42.12% |
| Non-Hispanic/Latinx | 428 | 56.69% |
| Gender | | |
| Female | 294 | 39.94% |
| Male | 449 | 59.47% |
| Other Gender | 10 | 1.32% |
| School Level | | |
| Secondary (MS/HS) | 329 | 43.58% |
| Tertiary (College) | 426 | 51.95% |
| FORPL Status | | |
| Eligible | 547 | 72.45% |
| Not Eligible | 208 | 25.55% |
| Parental Education Level | | |

**Table 1.** (continued)

| Demographic | Sample Size | Percentage |
|---|---|---|
| *Mother* | | |
| Some High School | 45 | 5.96% |
| High School Graduate | 131 | 17.35% |
| Some College | 175 | 23.18% |
| Bachelor's Degree | 257 | 34.04% |
| Graduate Degree | 124 | 16.42% |
| Other | 9 | 1.19% |
| *Father* | | |
| Some High School | 34 | 4.50% |
| High School Graduate | 110 | 14.57% |
| Some College | 168 | 22.25% |
| Bachelor's Degree | 266 | 35.23% |
| Graduate Degree | 141 | 18.68% |
| Other | 8 | 1.01% |

## Measures

### Demographic Questionnaire

A demographic questionnaire was distributed to all participants. This questionnaire included several items related to demographics including information on school attended, grade level, age, gender, race, ethnicity, subsidized lunch status, parental education level, and geographic location.

**The Abbreviated Math Anxiety Scale (AMAS; Hopko, Bare, & Hunt, 2003)**

The AMAS is a 9-item abbreviated version of the MARS-R (Plake & Parker, 1982), a revision of the MARS-A (Suinn & Edwards, 1982). On the AMAS, participants are asked to describe the extent to which they would be made anxious by several scenarios. Responses to items are given on a 5-point Likert-type scale, including the responses “not at all”, “a little”, “A fair amount”, “much”, or “very much”. The AMAS includes the two factors that made up the original MARS-A, learning math anxiety and math evaluation anxiety. Although it has been stated that the AMAS can be used on a high school sample (Ashcraft & Moore, 2009), the validation sample for the AMAS upon its creation was exclusively college undergraduate students. Reliability and validity analyses were conducted on the AMAS with a sample of 1,293 undergraduate students (n = 729 females, n = 510 males). The mean age of participants was 19.6 years. For internal consistency, the AMAS has a Cronbach’s Alpha score of .9. Additionally, it has a test-retest reliability score of .85. The AMAS has also demonstrated convergent and divergent validity. As expected, during the measure’s standardization, scores on the AMAS were convergent with that of the MARS-R, a reliable measure of math anxiety. Scores on the AMAS were divergent from math grades, number of math courses taken by raters, and measures of other types of anxiety including acute, chronic, and performance anxiety. Additionally, in the original validation study, the mean scores across ethnicity, in a White/European American vs. Minority dichotomous grouping, were largely equivalent. Sample items include “Having to use the tables in the back of a math book” and “being given a “pop” quiz in a math class”.

Although Hopko, Bare, and Hunt (2003) do not suggest cutoff scores for the measure, the mean of their primary sample was 21.1 with a standard deviation of 7.0. Notably, the mean was higher for female students (M = 21.9, SD = 6.9) than Male students (M = 19.5, SD = 6.9). Other math anxiety researchers have suggested differing cutoff criteria for the measure (Dykeman, 2017). For example, Maloney and Colleagues (2010) suggest participants scoring from below 20 have low math anxiety and those scoring above 30 have high math anxiety, Morsanyi, Busdraghi, and Primi (2014) suggest using a median split to classify low and high participants, and Ashcraft and Kirk (2001) suggest that, for math anxiety research generally, scores at least one standard deviation below the mean represent low math anxiety and scores at least one standard deviation above the mean represent high math anxiety.

The AMAS has been validated for high school students as well (Primi, 2014). However, the high school sample included 215 White/European American students (52% male) from a suburban area of Tuscany, Italy. The original AMAS measure was translated into Italian for Primi (2014), and has since been used in multiple studies with Italian samples (Buratta et al., 2019; Hill et al., 2016; Z. Wang et al., 2020). A forward translation method was used to translate the scale into Italian. Two non-professional translators worked independently and compared translations to assess the equivalence of the Italian version to the original English version. After the independent translations and their comparisons, a focus group of five individuals read and revised the Italian version several times before obtaining a final form (Primi et al., 2014). A confirmatory factor analysis yielded the expected two factors and the Cronbach's alphas fell within an acceptable range at .86 (CI .83-.88) for the learning math anxiety subscale, and .80 (CI

.83-.89) for the evaluation math anxiety subscale. Convergent validity was also found between the scale and test anxiety, as measured with the Test Anxiety Inventory (TAI; Spielberger et al., 1978). The AMAS has also been used with adolescents in UK (Devine et al., 2012) and China (Xie et al., 2019) through similar translation methods. While reliability and validity of such translated versions of the measure have been obtained, best practices in cross-cultural translation include five stages: forward translation, synthesis, back translation, expert committee review, and pretesting (Beaton et al., 2000). It will be important to take the lack of rigor in translation methods into account when evaluating the reliability and validity of the AMAS with the high school population sampled in this study, to ensure that the Standard American English version of the measure is also valid for use with high school populations.

**Mathematics Anxiety Scale-Revised (MAS-R; Bai, 2009)**

The MAS-R is a 14-item, two factor scale of math anxiety that is comprised of negative affect and positive affect towards mathematics. Each item on the MAS-R requested that participants agree or disagree with a given statement concerning affect towards mathematics. Sample items include “Mathematics makes me feel nervous” and “I find math interesting”. Responses to scale items range from 1 (strongly disagree) to 5 (strongly agree). For use with secondary school children, a sample of 647 7$^{th}$ and 8$^{th}$ grade students from pre-algebra and algebra classes were examined in a large public school district in the southeast United States. This sample was then split into a training sample (n=328) and a validation sample (n=343) to assess psychometric properties. In this sample, the MAS-R was found to have an internal consistency of .85, test-retest reliability of .71, and inter-factor correlation of .26. Additionally, it has been found to be convergent with math achievement (Bai, 2011). Psychometric support for the MAS-R has also been suggested for college undergraduate students (Bai, 2009). Cutoff scores for the MAS-R to determine high and low levels of math anxiety were not suggested by Bai (2009; 2011) nor were mean scores reported.

### Math Achievement

A subsample of one school was intended to include a question of math achievement for study participants. This was planned to be a school report measure in which data is requested from the school on the numeric grades of students who participate in the study in order to compare their math anxiety levels to their math achievement levels. However, due to barriers encountered with study recruitment, the math achievement measure was not included in the study.

## Procedures

### Pilot Sample

A pilot sample was collected through social media recruitment. For the pilot sample, administrators of 10 parent groups were contacted on Facebook. Researchers asked if a post could be made on their group pages with a short explanation of the study and a link to the parent consent form. Of these 10 groups, one group agreed to participate. A post was made on their page and 10 participants were obtained for the pilot sample. Participants were then recruited through a three-pronged approach: outreach to local schools, outreach to nonlocal schools, and outreach to parent groups on social media.

### Local School Outreach

Local school outreach first entailed reaching out to administrators of local high schools, through previously established relationships, and requesting a partnership with the research group to recruit students for the study. This outreach yielded a sample of convenience, as schools were selected via previously established relationships with researchers. However, due to the large sample size that is necessary for this study, it is most feasible to select schools through convenience sampling. Out of the five schools that were contacted during the local school recruitment phase, one school participated. This school helped recruit participants through distribution of consent forms via email and passing out physical copies of the consent form to students. In total, 60 consent forms were completed by parents, and 25 participants completed the study questionnaire. Four questionnaires were started and not completed. Outreach to local high schools also included 21 Charter schools, but no local charter schools agreed to participate during the data collection period.

**Nonlocal School Outreach**

Nonlocal school outreach entailed creating an email list of administrative contacts from school districts across the state of Pennsylvania as well as other states in the U.S., utilizing publicly available information from school district websites. To ensure sufficient numbers of Black/African American students to complete an MG-CFA with a Black/African American group, preference was given to schools that serve high numbers of Black/African American students in recruitment. This was completed through first intentionally searching for school districts in the United States with high Black/African American populations. Search engines utilized include Google and Bing. Several school lists were obtained through this method, including lists of high achieving schools with high Black/African American populations and lists of schools that are designed to serve Black/African American student. Schools were selected from these lists based on the percentage of Black/African American students served. Schools were rejected if their student population was less than 40% Black/African American. School demographics were obtained through free online services that provide demographic information about schools throughout the United States. This search yielded the first 25 schools. The additional 5 schools for the first wave of recruitment were obtained through randomly selecting states and searching for school districts in that state that serve high percentages of Black/African American students. Schools serving a population with less than 40% Black/African American students were rejected. Approximately one week after the first recruitment email was sent, a second reminder email was sent to those who had not responded. One week after the reminder email was sent, a third reminder email was sent to those who had not responded. As responses were obtained from schools, emails were sent describing next steps, which would include setting up a virtual meeting to discuss recruitment methods. During the first recruitment wave in the nonlocal recruitment phase, one school responded that they would like to participate. As the anticipated number of schools needed for recruitment is 15 schools, a second wave of recruitment emails was sent to another 30 schools, utilizing the same methods as the first. Recruitment continued at a rate of 30 schools per wave for 10 total waves. Of the 300 schools that were contacted, 2 schools participated in the recruitment process. The first school yielded 7 research subjects, and the second school yielded 0.

***Recruitment Plan with School Partners***

Once partnerships with local and nonlocal schools were established, next steps involved creating a plan for recruitment with each school. Each school had a different procedure for data collection, and such procedures were reviewed by researchers. All procedures required obtaining written permission from each school and a release of student contact information for recruitment. Consent from parents were gathered and questionnaires were distributed in accordance with each school's procedure.

For recruitment, teachers were asked to share the opportunity to participate in the study with their students in class, and researchers shared a presentation about math anxiety and the opportunity to participate in the study to attach to schools' newsletters and social media pages.

***Consent Procedures***

Consent procedures entailed sending consent forms to parents via email explaining the study and incentives and asking them to return the forms to the researcher's email address if they wish to have their children participate in the study. Two reminder emails were sent to parents with the consent form over two additional consecutive weeks. For local school recruitment, physical copies of the consent form were also passed out to students by researchers and teachers each Friday during the recruitment phase.

**Social Media Outreach**

Finally, the third recruitment method comprised of outreach on social media. Outreach included searching for relevant social media groups, and posting a request for study participation in these groups. Additionally, the parental consent form was posted broadly on numerous social media platforms. Parental consent was acquired through electronic distribution of consent forms to parents via social media messaging platforms and a request for an electronic signature returned to researchers. A total of 1706 parental consent forms were completed utilizing social media recruitment. Of those 1706, 805 study questionnaires were completed. However, several were removed during the data cleaning process, which is described in the Participants section.

### College/University Recruitment

After one year of recruitment, it was determined that the sample was not large enough to continue to recruit solely high school students, and college undergraduate students were added to the sample. College undergraduate students were recruited through posting the study questionnaire broadly on numerous social media platforms as well as reaching out to college administrators and collegiate organizations from colleges across the country. Preference was given to colleges that had high populations of Black/African American students, such as HBCUs and community colleges in areas with high Black/African American populations. Outreach was pursued with administrators from 64 colleges, and 34 collegiate organizations. Recruitment commenced for three weeks before an acceptable total sample was obtained. This sample is described in the participants section above.

### Incentives

Incentives were utilized in recruitment including sharing aggregate results of the study with schools and parents as well as a presentation related to math as professional development for teachers at the schools. Monetary incentives were utilized for study participants in the form of a raffle to win a $25 Amazon gift card for students who had participated in the study. The gift card amount was increased to $50 in May 2022, due to the low response rates garnered during research recruitment procedures. Consent forms and study measures were distributed to students by use of Qualtrics. Students completed and submitted one questionnaire that included consent, the demographic questionnaire, the AMAS, and the MAS-R through Qualtrics. This questionnaire also included a question asking students if they would be interested in being entered into the raffle to win an Amazon gift card. If they opted into the raffle, they were asked to include an email address to use for conducting the raffle and sending the gift card, should they win.

## Data Analysis

### Preliminary Analysis

A preliminary analysis of items examined descriptive statistics, normality, and correlations between items as well as reliability of scales. Multiple imputation (MI) procedures were used to account for missing data. Both Full Information Maximum Likelihood (FIML) and multiple imputation (MI) have been found to be more robust procedures for accounting for missing data than more traditional methods (Johnson & Young, 2011). Although FIML has been found to outperform MI in several respects (Allison, 2012), it has not yet been validated for use with maximum likelihood procedures for categorical data, as its use relies on the assumption of multivariate normality (Johnson & Young, 2011). As this study utilizes categorical data, it would not be appropriate to use FIML. However, simulation research suggests that the performance of MI is satisfactory with categorical data, particularly when utilizing the multiple imputation with chained equations (MICE) framework (Lang and Little, 2018; Van Buuren et. al., 2006). The MICE framework is based on Fully Conditional Specification, in which each incomplete variable is imputed by a separate model. This method can be used to impute mixes of continuous, binary, unordered categorical and ordered categorical data (Van Buuren et. al., 2006; Van Buuren & Groothuis-Oudshoorn, 2011). The MICE framework was utilized for the missing data analysis with the current study.

**Confirmatory Factor Analysis**

A confirmatory factor analysis was conducted on each measure to verify its factor structure. While previous research has suggested a two-factor model structure for both the AMAS (Hopko, 2003) and the MAS-R (Bai 2009, 2011), the existence of a bifactor structure and a higher order structure for the AMAS and the MAS-R were tested because scores on the dimensions on each of these scales have been found to be correlated with each other (Hopko, 2003; Bai 2009, 2011; Carey et al., 2017; Caviola et al., 2017). If the subscales are correlated with each other, they are likely also correlated with the total score on each of these measures. As this is the case, item level scores on each measure may also be correlated with the total scores on each measure. A regular correlated two-factor model was also tested for each measure. For the AMAS, this structure comprised two factors: learning math anxiety and math evaluation anxiety. For the MAS-R, this structure comprised two factors: positive affect and negative affect. Notably, the items for the positive affect measure are reverse coded to find the overall score as well as the positive affect score on the measure.

Post-hoc modifications were not explored in this study. Simulation research (Hutchinson, 1998) suggests that post-hoc modifications, utilizing modification indices, are consistent with sample sizes of 800 or more when model misspecification is low. However, if model misspecification is high, a sample size of at least 1200 is needed to accept post-hoc modifications (Hutchinson, 1998). As this study has a sample size of 755, post-hoc modifications may not be reliable even when model misspecification is low. It is also important to take theory and content validity into account when approaching post-hoc model modifications to avoid changing the model structure due to chance differences between the model and the data. In this study, when the post-hoc modification did not fit well with the theoretical basis for the measure, that modification was not pursued. Table 13 in Appendix B presents predictions of parameters that may be modified through post-hoc modification procedures with a theoretical rationale.

Reliability of each measure was also calculated in the preliminary analysis, utilizing Cronbach's alpha and McDonald's omega (McDonald, 2013). McDonald's omega has been found to have advantages over Cronbach's alpha as a test of reliability (Charles, 2020; Goodboy & Martin, 2020; Hayes & Coutts, 2020). Unlike Cronbach's alpha, for example, McDonald's omega does not assume unidimensionality of items on a scale, and it also allows factor loadings to vary (Hayes & Coutts, 2020). The benefits of Omega are useful for both the AMAS and the MAS-R, as they are both multidimensional scales, as they measure more than one factor. Additionally, to address Hypothesis 1, the reliability scores were compared across racial groups, and with the total sample, to determine the level of reliability for the total sample compared to the racial minority sample. The same procedure was applied to examine the reliability of each scale across gender groups and that of the total sample.

**Multi-Group CFA (MG-CFA)**

A multi-group confirmatory factor analysis (MG-CFA) were conducted across racial groups to determine the level of configural, metric, and scalar/threshold invariance across groups. MG-CFAs were only conducted with groups that had at least 200 participants in them. One MG-CFA was conducted with a White/European American student group and a Black/African American student group. An MG-CFA was unable to be conducted across gender groups within the Black/African American student group, due to limitations in sample size. MG-CFA was conducted using WLSMV estimation on R, as the data is categorical in nature and the sample size is over 200 (Liang & Yang, 2014). Model structure and fit across groups were examined using the following fit indices: chi-square Goodness of Fit, the root mean square error of approximation (RMSEA), the Standard Root Mean Square Residual (SRMR), the Comparative Fit Indices (CFI), and the Tucker Lewis Index (TLI). Criteria for model fit was derived from the structural equation modelling literature, which suggests an RMSEA <.06, an SRMR <.08, and a CFI and TLI >.95 (Hu & Bentler, 1999). The Chi-squared Goodness of Fit statistic was considered to hold the least weight when compared to the other fit indices, as it is affected by large sample sizes (Mulaik, 2009; Pendergast et al., 2017). Criteria for measurement invariance involved examining changes in model fit indices between models. First, structural invariance across groups was determined by model fit. When the same model displayed adequate fit across groups, the factor structure of math anxiety on the given scales was deemed invariant across groups. For the examination of structural invariance, no parameter constraints were applied across groups. To examine metric invariance, factor loadings were also inspected for equivalence of loadings on items on each factor across groups. In order for this inspection to occur, factor loadings were constrained across groups. At the metric invariance level, measures were considered non-invariant across groups when model fit worsened as constraints were applied. Finally, to examine scalar/threshold invariance, the equivalence of intercepts and thresholds between response options were evaluated across groups. For the examination of scalar/threshold invariance, both factor loadings and item means were set equal, or constrained, across groups. As with metric invariance, measures were considered non-invariant across groups when model fit worsened as constraints were applied (Mulaik, 2009; Pendergast et al., 2017). To determine if model fit worsened significantly, models of each measure for each invariance level were compared. The structural model of each measure was compared to the metric model of each measure, and the metric model of each measure was compared to the scalar model of each measure. Model fit differences were considered significant based on the change in the chi squared statistic as well as RMSEA and CFI. Cutoffs used include a change in CFI of $\geq$ -0.01, a change in RMSEA of $<$ 0.015, and a change in significance of the Chi squared (Chen, 2007).

The fourth level of measurement invariance, residual invariance, was not included in this analysis as research suggests that residual invariance may not have meaningful implications for measurement bias across groups (Brown, 2006; Chen, 2008).

When full invariance was not established for a measure, partial measurement invariance was examined for that measure. The examination of mean differences under the condition of partial invariance is controversial. Meredith (1993, 2006) in particular has argued that strict factorial invariance is required to avoid test bias in group comparisons. To avoid test bias in group comparisons, some researchers have argued that mean differences should never be explored if invariance is not met on the majority of the items on a measure (Cheung & Rensvold, 1998). Others have argued that, while partial invariance can be valid, the majority of factor loadings and intercepts should be invariant for the sake of reliability of the estimated latent means (Steenkamp & Baumgartner, 1998). However, there are researchers who have argued that at least half of the items on a scale should be invariant for comparisons across groups to be meaningful (Reise, Widaman & Pugh, 1993). Still, there is simulation research that suggests that mean comparisons can be conducted across groups if a measure meets criteria for partial invariance across those groups, with even only two indicators displaying invariance (Byrne et al., 1989; Shi et al., 2019). For example, using Monte Carlo simulations of multiple-group CFA analyses (MG-CFA), Shi and colleagues (2019) found that, while fitting fully invariant models or using reference items (RIs) with noninvariant loadings and/or intercepts biased mean comparisons across groups, fitting partially invariant models with only two items displaying invariance produced accurate cross-group comparisons of latent constructs, as long as one of those items was the reference, or marker, item (RI). However, it is important to note that Shi and colleagues (2019) also stress the importance of conducting a qualitative analysis of the item content along with the MG-CFA analysis to examine cross-cultural variance across indicators. Regardless of the criteria used to declare partial invariance for a measure, a review of measurement invariance research (Vanderberg & Lance, 2000) has found that about 50% of studies included in the review reported testing partial invariance. Schmitt and Kuljanin (2008) argue that most applied measurement invariance research will benefit from testing partial invariance, as it is common to find some non-invariance on measures across groups, and dropping items that display this invariance has implications for the reliability of those measures. In accordance with the argument of Schmitt and Kuljanin (2008) as well as several other researchers' perspectives on the utility of partial invariance (Bryne et. al. 1989; Steenkamp & Baumgartner, 1998; Reise, Widaman & Pugh, 1993; Shi et. al., 2019) this study included consideration of partial invariance when measures did not display full invariance.

For the analysis of partial invariance, post-hoc model modifications were examined in response to the cause of the variance at that level. For example, if one measure is variant at the structural level due to the presence of an item that does not load well onto its theorized factor with one group, that item will be dropped and the measure will be re-analyzed at that level without the offending item. Because this study has a research question inquiring about mean differences across groups, this study was open to conducting partial invariance testing utilizing evidence-based criteria that consider partial invariance with a lower number of indicators displaying invariance on each measure. Consistent with the simulation from Shi and colleagues (2019) as well as Byrne and colleagues (1989), measures were considered to have partial invariance across groups if the parameters (loadings and intercepts) of at least two indicators, or items, on that measure were equal across groups. Indicators that were variant across groups were released (Byrne et. al., 1989; Pirralha, 2020; Steenkamp & Baumgartner, 1998). However, in order to maintain the integrity of the scientific method, boundaries of post-hoc model modifications were pre-determined by theory. The items that had been predicted to have DIF, which can be found in Table 1 in Appendix A, were the only items that were allowed to be released, unless an item that had not been predicted in Table 1 had substantial differences across groups in one of the levels of analysis while utilizing the original models.

**Multiple Regression**

After measurement invariance was established, group differences by Race were examined. Following the work of Shi and colleagues (2019), this study conducted mean comparisons across groups when criteria for partial invariance was met on at least two items of a measure, including the reference item (RI), and a qualitative analysis of the items suggests similar functioning across groups. The scale with the higher levels of reliability and/or validity was used in the analysis of group differences. To compare group differences in levels of math anxiety experienced by students, a hierarchical linear regression was conducted with math anxiety as the outcome variable. Predicting variables included race and gender. FORPL, Mother Education Level, and Father Education Level were not included in the analyses, as prior research suggests that socioeconomic status is not related to math anxiety (Jazdzewski, 2011; Dowker, 2016; Flowerdew, 2021). However, as Ahmed (2018) found that individuals whose parental education level was lower tended to fall into the consistently high and increasing math anxiety trajectories, a hierarchical regression analysis that includes FORPL, Mother Education Level, and Father Education Level as predicting variables was included in the Appendix. Variables that predicted math anxiety were considered to have significant group differences. Effects coding was utilized for categorical predictor variables that had more than two categories. Gender was included first in the regression, as there is more consistent evidence for gender as a predictor of math anxiety than race.

# CHAPTER 4

# RESULTS

## Preliminary Analyses

### Missing Data Analysis

Of the 755 students in the total sample, there were not any respondents who completed every item on the study questionnaire. However, due to the nature of the questionnaire, there are several variables that may not be applicable to any single respondent, and several variables that required open-ended responses. When considering each variable, including open-ended variables, the total percentage of missing data was 11.81%. When removing the open-ended variables, with the exception of the "School" variable, the total percentage of missing data was 2.68%. This is relevant, as multiple imputation would not include open-ended variables.

An analysis of the pattern of missing data in R revealed that in 75 cases (9.93%), respondents skipped the "School" variable. This may have been due to a desire for respondents to remain confidential when they completed the questionnaire online. No other significant patterns emerged. Additionally, Little's Missing Completely at Random test was used in R to determine if data were missing at random. These results were not statistically significant, indicating that the data are missing at random ($X^2 = 3018$, df = 5089, $p > .05$) The results of the analysis of missing data suggest that multiple imputation can be used to impute missing variables, as the data that will be included in the analysis is missing at random (Jakobsen et al, 2017). Multiple Imputation was conducted utilizing the mice function in R.

## Descriptive Statistics

Descriptive statistics for the items on the MASR and AMAS scales are reported in Table 2. As can be seen in Table 2, each item for the MASR and the AMAS scales as well as the total score for each scale on the questionnaire meet the criteria for normality according to the cutoff scores recommended by Kim (2013), which indicate that an absolute skew value greater than 2 and an absolute kurtosis value greater than 7 would suggest a substantial departure from normality. Additionally, histograms of each scale item appear to be relatively normally distributed. Further, the mean of the total AMAS score (M = 27.51, SD = 7.04) is within one standard deviation of the mean of the validation sample completed by its developers (M = 21.1, SD = 7.0; Hopko, Bare, and Hunt, 2003), and within the cutoff scores suggested by Maloney and Colleagues (2010), who suggest that a score of 30 or higher on the AMAS is indicative of high math anxiety and a score of 20 or lower is indicative of low math anxiety. Table 19, which displays the total AMAS and MAS-R means per racial and gender group, can be found in the Appendix. Notably, the mean scores for each group are also within the cutoff scores suggested by Maloney and Colleagues (2010) for the AMAS.

**Table 2.** Descriptive Statistics for AMAS and MASR Scales

| Item | N | Min | Max | Mean | SD | Skew | Kurtosis |
|---|---|---|---|---|---|---|---|
| MASR1 | 755 | 1 | 5 | 2.40 | .96 | .62 | -.08 |
| MASR2 | 755 | 1 | 5 | 3.28 | .99 | -.19 | -.52 |
| MASR3 | 755 | 1 | 5 | 2.36 | .97 | .56 | -.01 |

Table 2. (continued)

| Item | N | Min | Max | Mean | SD | Skew | Kurtosis |
|---|---|---|---|---|---|---|---|
| MASR4 | 755 | 1 | 5 | 3.14 | 1.03 | -.08 | -.67 |
| MASR5 | 755 | 1 | 5 | 2.38 | .98 | .56 | -.07 |
| MASR6 | 755 | 1 | 5 | 3.35 | 1.02 | -.25 | -.47 |
| MASR7 | 755 | 1 | 5 | 3.19 | 1.07 | -.13 | -.73 |
| MASR8 | 755 | 1 | 5 | 3.67 | .99 | -.62 | .04 |
| MASR9 | 755 | 1 | 5 | 3.24 | 1.03 | -.17 | -.61 |
| MASR10 | 755 | 1 | 5 | 2.54 | 1.06 | .49 | -.36 |
| MASR11 | 755 | 1 | 5 | 3.16 | 1.06 | -.11 | -.70 |
| MASR12 | 755 | 1 | 5 | 2.51 | 1.07 | .42 | -.53 |
| MASR13 | 755 | 1 | 5 | 2.49 | 1.03 | .63 | -.08 |
| MASR14 | 755 | 1 | 5 | 3.26 | 1.02 | -.16 | -.59 |
| TotalMASR | 755 | 18 | 68 | 40.99 | 7.01 | -.25 | .98 |
| AMAS1 | 755 | 1 | 5 | 2.90 | 1.08 | -.03 | -.78 |
| AMAS2 | 755 | 1 | 5 | 3.01 | 1.04 | .03 | -.53 |
| AMAS3 | 755 | 1 | 5 | 3.14 | 1.11 | -.22 | -.68 |
| AMAS4 | 755 | 1 | 5 | 3.07 | 1.03 | -.09 | -.61 |
| AMAS5 | 755 | 1 | 5 | 3.09 | 1.08 | -.10 | -.73 |
| AMAS6 | 755 | 1 | 5 | 3.03 | 1.13 | -.14 | -.78 |
| AMAS7 | 755 | 1 | 5 | 2.94 | 1.12 | -.04 | -.74 |
| AMAS8 | 755 | 1 | 5 | 3.15 | 1.11 | -.16 | -.73 |
| AMAS9 | 755 | 1 | 5 | 3.15 | 1.08 | -.24 | -.61 |

Table 2. (continued)

| Item | N | Min | Max | Mean | SD | Skew | Kurtosis |
|---|---|---|---|---|---|---|---|
| TotalAMAS | 755 | 9 | 42 | 27.51 | 7.04 | -.45 | -.08 |

**Correlation Analyses**

Bivariate correlations between study variables are reported below in Table 3. Due to low numbers of participants across the race variable levels of Asian American, Native American, and Other Race as well as those within the Gender variable level Other Gender, the Race variable was restricted to Black/African American and White/European American, and the Gender variable was restricted to Male and Female for the correlation analysis. Additionally, given that the Mother Education and Father Education variables include Other as the sixth response, the parental education variables were restricted to responses 1-5. As can be seen in Table 3, the AMAS total score and the MAS-R total score had a small but statistically significant correlation (.25, $p < .01$). The AMAS also had such correlations with Mother Education Level (.10, $p < .01$), as well as strong statistically significant correlations with its subscales, learning math anxiety (.95, $p < .01$) and math evaluation anxiety (.91, $p < .01$), and a significant moderate correlation with the negative affect subscale on the MAS-R (.34, $p < .01$). The MAS-R had small but statistically significant correlations with Age (-.12, $p < .01$), Grade (-.22, $p < .05$), and Mother Education Level (-.08, $p < .05$), as well as the subscales of the AMAS, learning math anxiety (.21, $p < .01$) and math evaluation anxiety (.28, $p < .01$), and moderate to strong correlations with its own subscales positive affect (.68, $p < .01$) and negative affect (.79, $p < .01$). Other notable correlations include that of Learning Math Anxiety with

FORPL (-.13, $p < .01$), Mother Education level (.12, $p < .01$), and Father Education Level (.09, $p < .05$), and that of Positive Affect with Age (-.15, $p < .01$), Grade (-.13, $p < .01$), School Level (-.10, $p < .01$), and Mother Education Level (-.08, $p < .05$). No additional notable correlations were obtained.

**Table 3.** Bivariate Correlations between Study Variables

| | Age | Grade | SL | Race | Gender | FORPL | HIS | Med | FEd | AMAS | LMA | MEA | MASR | PA | NA |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Age | | | | | | | | | | | | | | | |
| Grade | .48** | | | | | | | | | | | | | | |
| SchoolLevel | .79** | .56** | | | | | | | | | | | | | |
| Race | -.23** | -.23** | -.36** | | | | | | | | | | | | |
| Gender | .00 | -.07 | -.07 | .22** | | | | | | | | | | | |
| FORPL | .05 | -.02 | .02 | .02 | .03 | | | | | | | | | | |
| Hispanic | .01 | .11** | .08* | -.25** | -.27** | .11* | | | | | | | | | |
| MotherEd | -.10** | .07 | -.12** | .19** | .13** | -.05 | -.06 | | | | | | | | |
| FatherEd | -.13** | -.06 | -.18** | .24** | .08* | .01 | -.13** | .74** | | | | | | | |
| AMAS | .05 | -.03 | .04 | .02 | .02 | -.07 | .01 | .10** | .07 | | | | | | |
| LMA | .08 | .02 | .05 | .04 | .03 | -.13** | -.02 | .12** | .09* | .95** | | | | | |
| MEA | .05 | .03 | .04 | .03 | .01 | -.01 | .03 | .06 | .02 | .91** | .73** | | | | |
| MASR | -.12** | -.22* | -.08 | -.02 | .02 | -.02 | -.01 | -.08* | -.04 | .25** | .21** | .28** | | | |
| PosAf | -.15** | -.13** | -.10** | .04 | .02 | .03 | -.03 | -.08* | -.01 | .01 | -.02 | .06 | .68** | | |
| NegAf | -.04 | -.01 | -.02 | -.06 | .00 | -.06 | .03 | -.05 | -.03 | .34** | .30** | .34** | .79** | .08* | |

Note. *p<.05, **p<.01. The positive affect subscale was reverse coded.

**Reliability Analyses**

The results of the reliability analyses, Cronbach's alpha, and McDonald's omega are reported in the table below for each scale. As shown in Table 4, the AMAS and MAS-R displayed acceptable reliability by the results of both Cronbach's alpha and McDonald's omega. Notably, the AMAS displayed a higher level of reliability across both analyses. In order to answer research question 1, the reliability of each scale with the total sample was compared to the reliability of each scale when racial and gender groups were split into subsamples. Statistical inferences were calculated between the reliability scores of the total sample and that of each subsample utilizing the software package cocron (Diedenhofen & Much, 2016) in R, which implements the methods described by Feldt et al. (1987) to calculate statistical inferences between reliability scores. Due to low sample sizes, the "Other Gender" and "Other Race" subsamples were not included in this analysis. The only reliability differences across subsamples that were statistically significant were the Cronbach's alpha and McDonald's omega scores between "Female" and the Total Sample with the AMAS (.85, .88, $p < .05$; .89, .86, $p < .05$), suggesting that the AMAS is more reliable with the Total Sample than among participants in the sample who identified as "Female", and the McDonald's Omega scores between the "Male" and the Total Sample with the AMAS (.89, .91, $p < .05$), suggesting that the AMAS is more reliable with the "Male" subsample than with the Total Sample.

**Table 4.** Cronbach's Alpha and McDonald's Omega for AMAS and MASR scales.

| Scale | Sample | Alpha | Omega |
|---|---|---|---|
| AMAS | Total | .88 | .89 |
| AMAS | Black/African American | .90 | .91 |
| AMAS | White/European American | .88 | .89 |
| AMAS | Asian American | .86 | .88 |
| AMAS | Native American | .89 | .92 |
| AMAS | Male | .90 | .91* |
| AMAS | Female | .85* | .86* |
| MASR | Total | .76 | .83 |
| MASR | Black/African American | .79 | .84 |
| MASR | White/European American | .75 | .84 |
| MASR | Asian American | .71 | .82 |
| MASR | Native American | .78 | .88 |
| MASR | Male | .77 | .84 |
| MASR | Female | .75 | .83 |

Note. *Statistically significant at p<.05, **Statistically significant at p <.01

**Confirmatory Factor Analyses**

Confirmatory factor analyses were performed across both the AMAS and the MAS-R. Bifactor and Higher Order models were attempted for each measure, but neither model would converge for either measure. Thus, a standard, correlated two-factor model was tested for each measure, which resulted in successful convergence. Results of the Confirmatory factor analyses are reported in Table 5 below.

As can be seen in the table below, the $\chi$ 2 was non-significant for the AMAS model, the CFI and TLI exceeded the recommended fit statistics ≥ .90, the RMSEA exceeded the recommended fit criterion < .06 and the SRMR exceeded the recommended fit < .08.

However, the $\chi$ 2 for the MAS-R model was statistically significant, and the CFI, TLI, and RMSEA were not within their recommended cut scores. SRMR was <.08 but was larger than that of the AMAS. Post-hoc modifications were suggested for the MAS-R, most of which included changes to item 8, which appeared to be a problematic item on the measure with this sample. However, as this was not a prediction based on theory, no post-hoc modifications were pursued.

**Table 5.** Fit Statistics for Standard Two Factor Models of the AMAS and MASR

| Scale | N | $\chi$ 2 | *Df* | P | CFI | TLI | RMSEA | SRMR |
|---|---|---|---|---|---|---|---|---|
| AMAS | 755 | 30.59 | 26 | .24 | .998 | .997 | .015 | .019 |
| MAS-R | 755 | 361.65 | 76 | .00 | .841 | .809 | .071 | .071 |

Note. $\chi$ 2 = chi-square. RMSEA= root mean square error of approximation. Df = degrees of freedom of the $\chi$ 2. CI = confidence interval. CFI = comparative fit index. TLI = Tucker Lewis index. Model pattern coefficients generated using WLSMV estimation method.

## Multi-Group CFA

Multi-group confirmatory factor analyses were conducted across race with the AMAS. The MG-CFA was not pursued for the MAS-R, as the CFA for the model did not display adequate fit. Due to sample size limitations, the Race variable was restricted to the Black/African American and the White/European American subgroups. Additionally, a multi-group confirmatory factor analysis was not able to be conducted between gender subgroups within the Black/African American racial group, due to sample size limitations. The Lavaan package in R was utilized to test configural, metric, and scalar invariance of the AMAS across race.

For the AMAS, the configural invariance test met criteria for acceptable fit: $\chi 2$ (df = 52) = 63.35, CFI = .994, TLI = .991, RMSEA = .026, SRMR = .027. When constraining the factor loadings to be equal across groups, the overall fit by itself was still acceptable: $\chi 2$ (df = 59) = 64.69, CFI = .997, TLI = .996, RMSEA = .017, SRMR = .030. The change in fit for the $\chi 2$ was 1.34 ($\Delta$ *df* = 7; p = .51), .003 for the CFI and .009 for the RMSEA. Based on the overall pattern, utilizing the cutoffs suggested by Chen (2007), we can determine that metric invariance was met for the AMAS. When constraining intercepts to be equal across groups, the fit indices for the AMAS were $\chi 2$ (df = 66) = 71.70, CFI = .997, TLI = .996, RMSEA = .016, SRMR = .032. The change in fit for the $\chi 2$ was 7.01 ($\Delta$ *df* = 7; p = .43), 0.00 for the CFI, and .001 for the RMSEA. Based on the overall pattern, as well as the cutoffs suggested by Chen (2007), we can determine that scalar invariance was met for the AMAS.

**Table 6.** Fit Statistics for the MG-CFA models of the AMAS by Race

| Model | N | χ 2 | *df* | *P* | Δ χ2 | Δ *df* | *P* | CFI | ΔCFI | TLI | RMSEA | ΔRMSEA | SRMR |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| AMAS | | | | | | | | | | | | | |
| Configural | 658 | 63.35 | 52 | .13 | N | N | N | .994 | N | .991 | .026 | N | .027 |
| Metric | 658 | 64.69 | 59 | .285 | 1.34 | 7 | .51 | .997 | .003 | .996 | .017 | .009 | .030 |
| Scalar | 658 | 71.70 | 66 | .295 | 7.01 | 7 | .43 | .997 | .000 | .996 | .016 | .001 | .032 |

Note. χ2 = chi-square; df = degrees of freedom. Δ χ 2 = change in the root mean square error of approximation. Δdf = change in the degrees of freedom. RMSEA= root mean square error of approximation. CFI = comparative fit index. ΔCFI = change in the comparative fit index. TLI = Tucker Lewis index. N = not applicable. Model pattern coefficients generated using WLSMV estimation method.

As the study sample included both secondary and tertiary level students, an MG-CFA was conducted for the AMAS across the school level variable, to determine measurement invariance across school level. For the AMAS, the configural invariance test met criteria for acceptable fit: $\chi$ 2 (df = 52) =69.34, CFI = .991, TLI = .988, RMSEA = .030, SRMR = .027. When constraining the factor loadings to be equal across groups, the overall fit by itself was still acceptable: $\chi$ 2 (df = 59) = 62.82, CFI = .998, TLI = .998, RMSEA = .013, SRMR = .031. The change in fit for the $\chi$ 2 was 2.88 ($\Delta$ *df* = 7; p = .90), .010 for the CFI and .017 for the RMSEA. Based on the overall pattern, as well as cutoffs suggested by Chen (2007), we can determine that metric invariance was met for the AMAS. When constraining intercepts to be equal across groups, the fit indices for the AMAS were $\chi$ 2 (df = 66) = 81.80, CFI = .992, TLI = .991, RMSEA = .025, SRMR = .034. The change in fit for the $\chi$ 2 was 18.98 ($\Delta$ *df* = 7; p = .01), 0.10 for the CFI, and .017 for the RMSEA. Based on the overall pattern, and the cutoffs recommended by Chen (2007), we can determine that scalar invariance was met for the AMAS. However, it is important to consider that the change in fit for the $\chi$ 2 from metric to scalar was significant, suggesting that the fit of the model may have worsened significantly from metric to scalar for the AMAS across School Level. When further investigating the results of the scalar invariance analysis, we can conclude that this may be related to slightly differing intercepts obtained across school level groups. For example, six items had slightly higher intercepts for the secondary group than the tertiary group, and 3 items had slightly lower intercepts for the secondary group than the tertiary group. However, the highest difference in intercepts across groups was for item 9 at 0.292, which was higher for secondary students than tertiary students. Additionally, the change in CFI and

RMSEA were not significant from the metric to the scalar model, and these fit statistics have more weight than the $\chi$ 2 fit statistic. Table 18, which can be found in the Appendix, displays intercepts for the AMAS across school level.

**Table 7.** Fit Statistics for the MG-CFA models of the AMAS by School Level

| Model | N | $\chi 2$ | *df* | *P* | $\Delta \chi 2$ | $\Delta$ *df* | *P* | CFI | ΔCFI | TLI | RMSEA | ΔRMSEA | SRMR |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| AMAS | | | | | | | | | | | | | |
| Configural | 755 | 69.34 | 52 | .05 | N | N | N | .991 | N | .988 | .030 | N | .027 |
| Metric | 755 | 62.82 | 59 | .34 | 6.52 | 7 | .80 | .998 | .010 | .998 | .013 | .017 | .030 |
| Scalar | 755 | 81.80 | 66 | .091 | 18.98 | 7 | .01 | .992 | .006 | .991 | .025 | .012 | .034 |

Note. $\chi 2$ = chi-square; df = degrees of freedom. $\Delta \chi 2$ = change in the root mean square error of approximation. Δdf = change in the degrees of freedom. RMSEA= root mean square error of approximation. CFI = comparative fit index. ΔCFI = change in the comparative fit index. TLI = Tucker Lewis index. N = not applicable. Model pattern coefficients generated using WLSMV estimation method.

## Multiple Regression

For the multiple regression analysis, a hierarchical design was employed which used two models. Each model predicted Math Anxiety through use of the two subscales of the AMAS: Learning Math Anxiety and Math Evaluation Anxiety. While the total math anxiety score was originally planned to serve as the predicted variable, results of the CFA analyses suggest that the best model for the AMAS is a simple, two factor structure as opposed to a bifactor or higher order structure. As such, it was most appropriate to conduct two hierarchical linear regression analyses: one with learning math anxiety as the predicted variable, and one with math evaluation anxiety as the predicted variable. For each analysis, Model 1 only included one predicting variable: gender. Model 2 included gender and race as predicting variables. The order of the variables was decided based on previous literature, which more consistently finds gender associated with math anxiety, when compared to race, in which there are discrepancies in findings with respect to associations with Math Anxiety. Additionally, due to limitations of sample size, the gender variable was restricted to those who identified as "Male" and "Female", and the race variable was restricted to those who identified as "Black/African American" or "White/European American". The reference group for the gender variable was "Male", and the reference group for the race variable was "Black/African American". As can be seen in Table 8-11, neither gender nor race were predictive of Learning Math Anxiety or Math Evaluation Anxiety. This suggests that there are not significant differences across race and gender in level of Learning Math Anxiety or Math Evaluation Anxiety. Notably, two additional hierarchical regression analyses were conducted with all three levels of gender and all five levels of race, and two additional hierarchical regression analyses were conducted which included FORPL, Mother Education Level, and Father Education Level. No meaningful differences in results were obtained from these analyses.

**Table 8.** Hierarchical Linear Regression Model 1 with Gender as Predictor of Learning Math Anxiety

| Coefficient | Estimate | SE | p-value |
|---|---|---|---|
| Intercept | 14.7770 | 0.5117 | <2e-16** |
| Gender | 0.2429 | 0.3525 | 0.491 |

*Note:* $F(1,637) = 0.4748$, $p = .491$, $R^2 = 0.0007448$, $R^2_{adj} = -0.0008239$, $**p < .01$

**Table 9.** Hierarchical Linear Regression Model 2 with Gender and Race as Predictors of Learning Math Anxiety

| Coefficient | Estimate | SE | p-value |
|---|---|---|---|
| Intercept | 14.2975 | 0.7298 | <2e-16** |
| Gender | 0.1687 | 0.3616 | 0.641 |
| Race | 0.3454 | 0.3748 | 0.357 |

*Note:* $F(2,636) = 0.662$, $p = 0.5162$, $R^2 = 0.002078$, $R^2_{adj} = -0.001061$, $**p < .01$

**Table 10.** Hierarchical Linear Regression Model 1 with Gender and Race as Predictors of Math Evaluation Anxiety

| Coefficient | Estimate | SE | p-value |
|---|---|---|---|
| Intercept | 12.1076 | 0.3944 | <2e-16** |
| Gender | 0.1009 | 0.2716 | 0.71 |

*Note:* $F(1,637) = 0.138$, $p = 0.7104$, $R^2 = 0.0002165$, $R^2_{adj} = -0.001353$, $**p < .01$

**Table 11.** Hierarchical Linear Regression Model 2 with Gender and Race as Predictors of Math Evaluation Anxiety

| Coefficient | Estimate | SE | p-value |
|---|---|---|---|
| Intercept | 12.4276 | 0.5625 | <2e-16** |
| Gender | 0.1504 | 0.2787 | 0.590 |
| Race | -0.2305 | 0.2889 | 0.425 |

*Note:* $F(2,636) = 0.3873$, $p = 0.679$, $R^2 = 0.001217$, $R^2_{adj} = -0.001924$, $^{**}p < .01$

# CHAPTER 5

# DISCUSSION

Despite good grades in mathematics courses, academically talented women and individuals who identify with racially marginalized groups often do not pursue careers in STEM (Kerr & Kurpius, 2004; McCormick, 1998), and there is a lack of representation of women and individuals who identify with racially marginalized groups in STEM fields (National Science Foundation, 2019). The lack of representation of women and individuals from racially marginalized groups in STEM fields is problematic for a variety of reasons, including the lack of equity in career options across marginalized groups, and the absence of diversity in the development of STEM products. As such, attrition and retention of women and racially marginalized individuals in STEM programs at the college level as well as in mathematics courses at the secondary level is crucial.

One way to approach this problem is to address students' math anxiety. Math anxiety is a social-emotional factor that is associated with decreased levels of interest and retainment in mathematics courses (Ashcraft & Moore, 2009; Ramirez et al., 2018). A variety of individual research studies and meta-analyses have suggested that math anxiety levels are higher for girls and women (Buratta et al., 2019; Hembree, 1990; Hill et al., 2016; Luttenberger et al., 2018; Primi et al., 2014), but more mixed results have been found for math anxiety levels across racial groups. While some studies suggest that individuals from marginalized groups are more math anxious than White/European American individuals (Hembree, 1990; Ahmed, 2018), other studies do not show such differences (Fuson, 2007; Merrit, 2011; Carrol, 2010; Clark, 2004). However, math anxiety measures have historically not included representative numbers of non-

White/European American individuals in their standardization samples (J. R. Young & Young, 2015, 2016). It is important to accurately assess math anxiety in students in order to provide appropriate interventions, and to understand the role that math anxiety may play in larger disparities, such as those described previously. As such, this study set out to investigate the psychometric properties of the AMAS and MAS-R across race and gender groups.

## Summary of Results

Findings from this study suggest that the differences in reliability of the AMAS and MAS-R are not statistically significant across subsamples of race and gender, with the exception of the Female and Male subsamples and the Total Sample of the AMAS. Results of the reliability analysis suggest that the AMAS is more reliable with the Total Sample than with the Female subsample, and with the Male subsample than with the Total Sample. However, it is important to consider that the difference in reliability scores of the AMAS measure across the Total Sample and the Female sample was only .03, and across the Total Sample and the Male subsample was only .02, and that each score was above .8, suggesting that, while the difference between reliability scores may be statistically significant, it may not be practically significant. Results of the reliability analyses suggests that these measures are all sufficiently reliable across race and gender groups, indicating that the reliability of math anxiety measures may not be a significant concern when considering selection of interventions for individuals of racially marginalized backgrounds and girls.

In terms of validity, this study found that the AMAS is non-invariant across White/European American and Black/African American racial groups as well as secondary level and tertiary level students. The non-invariance of the measure across White/European American and Black/African American racial groups as well as secondary and tertiary level students suggests that measurement bias is not a concern when comparing math anxiety levels across these groups. In addition, the measurement non-invariance of the AMAS across these groups suggests that the AMAS is a valid measure to use as a screening measure in the determination of math interventions for Black/African American and White/European American students as well as secondary and tertiary level students.

Further, results of the hierarchical regression analysis indicate that race and gender are not predictive of learning math anxiety or math evaluation anxiety. This suggests that there are not significant differences in learning math anxiety or math evaluation anxiety between White/European American and Black/African American students or between Males and Females as sampled in this study. Taken together, these results suggest that math anxiety may not be a significant factor in the development of the differences in attrition and retention rates in STEM programs.

## Implications

### Consistency of Math Anxiety Measures Within and Across Groups

This study has multiple implications for many different stakeholders. Results of the reliability analysis indicate that the AMAS and MAS-R measures are not more reliable for students who identify as Black/African American, Asian American, or Native American when the samples being analyzed include only members of their own racial group. This is discrepant with previous research that has found that the MARS-A is more reliable with racially homogenous samples of Black/African American students (Young

& Young, 2015, 2016). However, it is important to consider sample size when comparing reliability scores across groups. For example, the total sample had 755 participants, while the Black/African American sample had only 208 participants, the Asian American sample had only 53 participants, and the Native American sample had only 36 participants. The size difference may have contributed to the lack of significant differences found in reliability of math anxiety measures across these groups. Future research might seek to further examine the reliability of math anxiety measures across race with more equal sample sizes across racial groups. These results further highlight the limitations in research findings due to sample sizes that do not contain adequate representation of individuals who identify as members of marginalized groups. Further, due to challenges with obtaining adequate sample sizes across racial groups for group comparison, future research might shift from a focus on group comparisons to a focus on within group similarities and differences with experiences of math anxiety.

**Factor Structures of Math Anxiety Measures**

Results of the Confirmatory Factor Analyses suggest that the AMAS has stronger evidence for its validity than the MAS-R for both secondary and tertiary school students, and thus the AMAS might be chosen over the MAS-R when researchers or school personnel are aiming to measure students' math anxiety. For example, the theorized 2-factor structure of the AMAS displayed adequate fit for the data in CFA across the whole sample as well as in separate analyses of both the secondary and tertiary level students. This suggests that the AMAS is valid for use with both age ranges. Additionally, this factor structure corresponds with the factor structure obtained in previous research that has been completed utilizing this measure across tertiary students in the United States

(Hopko, Bare, & Hunt, 2003), and secondary students in Italy (Primi, 2014), displaying the consistency of the measure across these populations. Further, results of the confirmatory factor analysis suggest that the AMAS has the capacity to measure both learning math anxiety and math evaluation anxiety for students, thus informing researchers and school personnel on two important and distinct aspects of math anxiety.

In contrast, the MAS-R displayed structural inconsistencies when compared to previous research. The theoretical 2-factor structure of the MAS-R, which includes the factors of “positive affect” and “negative affect” did not fit adequately with the study’s data. While previous research has found evidence for such a 2-factor structure utilizing EFA with undergraduate students (Bai, 2009), and both EFA and CFA with secondary level students (Bai, 2011), this study was unable to replicate these results in CFA with a sample containing both secondary and tertiary students, or when splitting the sample into secondary level and tertiary level subsamples. Notably, the sample size of Bai (2009) was only 78, and the sample from Bai (2011) contained 647 $7^{th}$ and $8^{th}$ grade students. The current study contained a sample with a majority of students at the college (N = 426) or high school (N = 289) level, and only 39 students at the middle school level. This age and school level difference may have contributed to the differences in model fit obtained in the current study compared to that obtained by Bai (2009, 2011).

Results of the confirmatory factor analysis suggest that researchers might explore alternative factor structures of the MAS-R, and/or make adjustments to item 8, “I find math challenging”, to improve the validity of the measure. The CFA analysis, as well as post hoc analyses, suggest that item 8 did not load well onto either factor of the measure with the current study’s population. It is possible that this item does not load well onto either “positive affect” or “negative affect” because individuals who find math challenging may feel positively or negatively about that challenge. Some individuals enjoy a challenge while others may be more frustrated and/or overwhelmed by that challenge. As such, the validity of the measure may be improved if this item was adjusted or removed.

Further, as results of the CFA on the AMAS and MAS-R in the current study suggest a two-factor model rather than a bifactor or higher order model, it is likely that there is not an overall construct of “math anxiety” that is being measured by the AMAS or MAS-R. As the AMAS is a shortened version of the MARS-A, which was utilized by Merritt (2011), Carroll (2010), and Clark (2004), utilizing the total score of the MARS-A may not yield consistent results. Instead, researchers might look to using the subscales of the AMAS in the future when studying math anxiety.

**Measurement Bias Across Race and School Level**

Results of the multigroup confirmatory factor analysis suggest that the AMAS is invariant across Black/African American and White/European American students, which indicates that the measure can be effectively used to compare math anxiety across these two racial groups. This has implications for researchers who might choose the AMAS as a measure of math anxiety when comparing White/European American and Black/African American students. Additionally, this illustrates the measure’s validity for use with Black/African American students, suggesting that researchers conducting research with this population may wish to use the AMAS. The invariance of the AMAS across these racial groups also has implications for school personnel of schools that have high percentages of Black/African American students, as it would be important for those individuals to select a measure of math anxiety that is valid for their students to aid in the selection of appropriate interventions.

## The Relationships between Gender, Race, and Math Anxiety

The results of the hierarchical regression analysis, which indicates that gender and race were not significant predictors of learning math anxiety or math evaluation anxiety, are not invalid due to non-invariance of the AMAS across racial groups. These results indicate that there are not significant differences in learning math anxiety or math evaluation anxiety across gender or race in secondary and tertiary school students. Notably, this finding is discrepant with previous research that has found differences in math anxiety across gender (Dowker, 2016; Hembree, 1990; Griggs et al., 2013; Yüksel-Şahin, 2008; Huang et. al., 2019; Hill et. al., 2016; Luttenberger et. al., 2018; Buratta et. al., 2019; Primi et. al., 2014; Else-Quest et. al., 2010), and race (Hembree, 1990; Ahmed, 2018). Further, these results suggest that math anxiety may not be a significant source of the disparities in academic achievement between White/European American individuals and individuals from marginalized racial groups. As such, school personnel might explore alternative sources of such disparities and focus their intervention efforts in other areas in order to mitigate the achievement gap.

### *Discrepancies in Math Anxiety Differences across Gender*

A variety of research studies conducted with samples from several different countries in Europe, from China, and from the United States suggest that math anxiety is higher for girls than for boys (Burrata 2019; Hill 2016; Primi 2014; Else-Quest 2010; Devine 2012; Frenzel 2007; Jain & Dowson 2009; Kvedore 2012; Lou, Wang, & Lou 2009). Notably, research has suggested that gender differences in math anxiety tend to be higher in countries that have comparatively low levels of math anxiety (Luttenberger, 2018). If this is true, the opposite, that gender differences in math anxiety tend to be lower in countries that have high levels of math anxiety, may also be true. The research that has suggested a lack of gender differences in math anxiety levels took place in various high schools throughout the United States (Ahmed, 2018; Tapia & Marsh, 2004), and middle schools in Turkey (Birgin 2010), and Finland (Kyattala & Bjorn 2014). It may be the case that there are higher levels of math anxiety in these countries, such that gender differences do not emerge.

However, Birgin (2010) notes that there have been other Turkish studies that have suggested that there are gender differences in math anxiety levels in Turkey (Bindak, 2005). Birgin (2010) suggests that the differences in findings within Turkish studies may be due to overall differences in math anxiety levels across regions. For example, Bindak (2005) sampled from a region in Turkey that had higher levels of math anxiety than the region in which Birgin (2010) sampled from as well as higher levels of math anxiety compared to typical math anxiety levels across Turkey. This difference in math anxiety levels across regions may serve as one explanation for why Bindak (2005) found gender differences in math anxiety levels while Birgin (2010) did not. Further, as this is the case, math anxiety differences across gender may be found not just across countries, but also regions within countries. Overall, this suggests that it is important to consider the cultural/environmental factors of the area in which data is being collected on math anxiety when interpreting gender differences in research.

Additionally, we might consider overall levels of math anxiety in the current study compared to other studies. The mean score on the total AMAS ($M = 27.46$, $SD =$

7.05) in the current study was higher than that of its validation sample (M = 21.1, SD = 7; Hopko, 2003). However, it is important to consider that the difference in mean scores between this study and the validation sample of the AMAS was within one standard deviation as calculated by both studies. Further, the mean for the total AMAS score as found in this study was not at or above 30, the suggested cutoff criteria for high math anxiety by other researchers (Maloney et al. 2010).

We might also consider effect sizes as well as temporal factors between studies that have found differential results for math anxiety levels across gender. Notably, there have been small effect sizes found in previous research that has found higher math anxiety for girls and women when compared to boys and men (Hembree, 1990, d = .19; Hill, 2016, d = .28), and small effect sizes are more difficult to replicate. Further, Hembree was a meta-analysis that was conducted in the United States in 1990, which included studies from up to 20 years before its release, which was 32 years before data was collected for the current study. It is possible that the higher math anxiety levels that were found for girls and women when compared to boys and girls at that time have shifted due to generational changes in the conceptualization of math as a white, male domain.

An additional consideration is that, when breaking down math anxiety by factor, some studies find that, while women tend to have more math evaluation anxiety/test anxiety, men tend to have more learning math anxiety/numerical anxiety (Luttenberger 2018). The current study found that gender was not predictive of either learning math anxiety or math evaluation anxiety at the tertiary level, suggesting that women and men do not differ significantly in either component of math anxiety. Future research might continue to examine differences in math anxiety across gender.

### ***Discrepancies in Math Anxiety Differences across Race***

***Methodology.*** The current study utilized hierarchical regression to determine the predictability of race for math evaluation anxiety and learning math anxiety. This was a different method than previous research. The methodology employed by Ahmed (2018) examined math anxiety through a longitudinal approach and utilized odds ratios to determine the differences in trajectories of math anxiety across time for different racial groups. The methodology utilized by Ahmed, as well as that utilized for the current study, was less dependent on sample sizes than the approach utilized by Merritt (2011), Carroll (2010), and Clark (2004), all of whom used an analysis of variance test to compare mean differences in total math anxiety across racial groups. As researchers tend to have challenges in recruitment of non-white/European American individuals, it may be beneficial to adopt methodologies that do not rely on large sample sizes for group comparisons. Further, these divergent research methods may have implications for the conclusions of each study. For example, it may be the case that there are not significant mean differences in math anxiety across race as measured by the MARS-A, and that race is not predictive of learning math anxiety or math evaluation anxiety as measured by the AMAS. However, it may also be true that Black/African American students tend to report consistently high math anxiety more often than White/European American students across time.

***Demographics.*** In terms of demographics, the current study and the studies cited above included different racial groupings and school levels in their analyses. This study grouped “secondary level” students as those who reported their school level as middle or high school, and “tertiary level” students as those who reported their school level as college. This study included five racial identity options: White/European American, Black/African American, Asian American, Native American, and “Other” with the option of writing in one’s racial identity. The questionnaire requested that individuals who identified with multiple races indicate “Other” and write in each race they identify with. The study did not include “Hispanic” as a racial identity option, but instead included an ethnicity item that requested participants indicate if they identified as Hispanic. Due to sample size limitations, this study only included “Black/African American” and “White/European American” when examining Race as a predictor of learning math anxiety and math evaluation anxiety.

Previous research that examined math anxiety levels across race had alternative racial identity groupings, and differing school levels included in their sample. For example, Hembree (1990), was a meta-analysis that included a variety of racial groups such as White/European American, Black/African American, Asian American, and Hispanic/Latino American in tertiary school students. Hembree (1990) found statistically significant differences in math anxiety only for the Hispanic/Latino group, suggesting that Hispanic individuals are more math anxious than White/European American individuals, as well as other ethnic groups. In contrast, Carroll (2010) compared White/European American adolescents to a racially heterogenous “minority” sample, Clark (2004) compared Hispanic/Latino American and Black/African American adolescents, and Merritt (2011) compared White/European American and Black/African American adolescents. Carroll (2010), Clark (2004), and Merritt (2011) found no statistically significant differences in level of math anxiety across race. Fuson (2007) also found no statistically significant differences in math anxiety between a college sample of White/European American students and a "minority" ethnic group college sample, which consisted of Black/African American, Asian American, and Native American students as well as an "other" ethnic group category of students.

Taken together, the results of these studies along with that of the current study suggest that more accurate and consistent results can be found when racially marginalized individuals are not grouped together in one sample and compared to White/European American individuals, but instead grouped separately by their self-identified racial group. Further, it may be beneficial to further explore the measurement invariance of the AMAS as well as the manifestation of math anxiety with secondary and tertiary level students who identify as Hispanic/Latino, as previous research has found differing results when comparing math anxiety levels of Hispanic/Latino individuals with those of other races, and the current study did not focus on these individuals.

### *Alternative Factors to Consider for Intervention*

It will also be important for school personnel to explore alternative interventions to support racially marginalized students, and address disparities in math achievement for racially marginalized students, as math anxiety is not the only barrier to success in higher level mathematics courses. For example, research has suggested that three intersecting identities are important for math success for Black/African American students: racial identity, disciplinary identity, and academic identity (Varelas, Martin, and Kene, 2013). School personnel might focus intervention efforts on the fostering of such identities for Black/African American students.

One way to achieve this might be through the use of socializing agents, such as parents, teachers, and the curriculum (Young, Young, and Capraro, 2017). Through the interaction of such socializing agents, school interventions might focus on creating a culturally responsive, strengths-based, and academically affirming environment to support the Black/African American students (Young, Young, and Capraro, 2017). Relatedly, research has suggested that Black/African American girls' math achievement is strengthened when teachers exhibit affirmative interpersonal relationships with them as well as create a sense of belonging for them in the classroom (Booker & Lim, 2018). Given the information provided by Booker and Lim (2018) as well as Young, Young, and Capraro (2017), schools might seek to adjust instruction to be more culturally responsive for Black/African American students, foster affirming interpersonal relationships between teachers and students, foster a sense of belonging for Black/African American students, and increase communication with the family members of Black/African American students in order to maintain interventions at home as well as at school in order to continue to foster the identities necessary for math success.

School personnel might also consider external factors such as racial biases in discipline practices when developing tiered interventions to support racially marginalized students, as racial biases in discipline, such as overrepresentation of Black/African American high school students in both in and out of school suspensions, have been identified as barriers to success in mathematics and overall STEM readiness for racially marginalized students (Ibrahim & Johnson, 2020). As such, school personnel might wish to focus their efforts on reviewing the impact of racial biases in their own school and choosing specific areas to target based on their findings. For example, a school might find that harsher punishments have been utilized for racially marginalized students in their district when compared to White/European American students for similar disciplinary infractions. As a result of this finding, a school might add policies that require examination of racial bias when making disciplinary decisions for students.

## Limitations

There were several limitations to the current study. First, the sample size was limited in that only two racial groups could be examined by MG-CFA. Future research might examine MG-CFA across additional racial groups as well as gender within race. Relatedly, sample size in general was a limitation of the study as 200 per group is the minimum suggested by many researchers for multi-group confirmatory factor analysis. Larger sample sizes per group might increase the generalizability of the research results. Relatedly, a power analysis was not conducted to determine that sample sizes were adequate to calculate study results. Future research might also ensure that a power analysis is conducted, especially when sample sizes are low.

Further, an adequate sample size was unable to be obtained for students between the ages of 13 and 18, and thus college students needed to be added to the sample. This addition compromised the study's ability to obtain results pertaining to math anxiety exclusively for adolescent students. Future research might include improvements on research methodology that allow for the focus to remain on adolescent students exclusively. Much of this challenge was due to difficulty obtaining parental consent for adolescents to participate in the study. Future research might adjust the methodology to improve on this challenge, such as the utilization of an opt out form.

This study was also limited in that the tertiary level sample was restricted to college students. Students who have the greatest math anxiety may be less likely to enroll in college. It is possible that differences in math anxiety levels across race and gender are more apparent when the sample includes post-secondary individuals who did not enroll in college.

The study further limited in that the majority of the data had to be collected remotely. Although bot detection software was utilized and criteria for inclusion in the study was stringent, it is still possible that bots made their way into the study or that individuals entered the study multiple times. Relatedly, the questionnaire was self-report, and thus researchers had to rely on participants to report on their demographic variables and math anxiety level accurately and truthfully. Similarly, many respondents (N = 167) did not report the school that they were attending, possibly due to the study data collection occurring primarily online. Due to this lack of response on school attended coupled with the low numbers of students in the sample attending the same school, the study did not include nesting of data. It is possible that many of the individuals who did not report the school that they attended were in fact attending the same schools, and thus nesting data would have been helpful when determining results and conclusions.

Further, racial identity was a significant variable in this study, and racial identity is complex for many individuals. It is possible that, by including a set number of racial identity options, the study did not capture the true scope of racial identity for the individuals participating. Due to this, results may not fully capture the level and complexity of the manifestation of math anxiety for individuals and groups based on the racial identities they reported. One example of this is the inclusion of Hispanic/Latino as a separate variable, under ethnicity. Many individuals consider Hispanic or Latino to be their racial identity. Indeed, many participants indicated “Other” for racial identity, and typed in “Hispanic” as their qualitative response to the racial identity question. As this is the case, a “Hispanic” racial group was not fully captured in either the racial identity or the ethnicity variable. Future research might include a qualitative component to their methods in order to address this limitation.

Similarly, an adequate sample was not obtained to include individuals who identify with a gender other than “Male” or “Female”. Low sample sizes for individuals with nonbinary gender identities limited the study’s ability to investigate reliability and validity of the math anxiety measures with nonbinary individuals. Future research might seek to utilize methods that are more appropriate for low sample sizes in order to obtain meaningful findings with this population. Qualitative methodologies might be explored for research with nonbinary individuals as well.

Finally, this study was limited in that it was not able to collect data on math achievement to examine associations between math anxiety and math achievement across race and gender. Previous research has suggested that higher levels of math anxiety are associated with lower levels of math achievement, and that gender can be a moderating variable within this relationship, such that math anxiety tends to have a stronger affect on math achievement for girls than boys. It is important to consider that math achievement was not included in this investigation, as findings about differences in math anxiety level cannot be examined in tandem with math achievement to determine if there are mediating and/or moderating variables in the association between math achievement and math anxiety with the current sample. Such relationships might have implications for math achievement among racially marginalized individuals, as race did not emerge as a predicting variable for learning math anxiety or math evaluation anxiety. As with some studies that examined the manifestation of math anxiety and achievement across gender, it may be that race is a moderating variable in the relationship between math anxiety and achievement, even when math anxiety levels are not significantly different across racial groups. Future research might aim to future examine the relationships between math anxiety, math achievement, race, and gender.

## Future Directions

Researchers might continue to explore math anxiety in specific racial groups in order to investigate the manifestation of math anxiety across racially diverse groups of students, to further aid in the selection of appropriate interventions for math anxious students. This research might also include an examination of the manifestation of math anxiety across gender groups within specific racial groups, to determine how race and gender interact in the manifestation of math anxiety. This research might include studies that employ mixed methods design in order to capture qualitative data to account for the complexity of racial and gender identities and the experiences that individuals who identify as members of marginalized racial and/or gender groups face which may shape their experiences with math and anxiety. Relatedly, future research might continue to examine levels and experience of math anxiety across gender specifically, as this study found results discrepant with previous research looking at math anxiety levels across gender. Such studies may also benefit from the inclusion of qualitative methods to better understand the experience of math anxiety as well as potential solutions. Further, future research may aim to explore and compare math, anxiety, and math anxiety specific interventions to investigate how measurement of math anxiety can inform the selection of appropriate interventions for students.

Future research might also wish to further explore the factor structures of the MAS-R and AMAS. First, the bifactor and higher order models of both the AMAS and MAS-R were unable to converge in the current study. This suggests the possibility that a higher order factor of math anxiety does not exist for these measures, and instead math anxiety manifests as either math evaluation anxiety or math learning anxiety as measured by the AMAS and/or positive affect or negative affect as measured by the MAS-R. However, significant, strong, positive correlations were obtained between the total score on each of these measures and each of their subscales, as well as between the subscales themselves. This suggests that the convergence problems may have been due to the high number of parameters of each model rather than the existence of a higher order factor for each model. As such, future researchers might further explore a bifactor and higher order model for the AMAS and MAS-R with a larger sample size, as increasing sample size tends to increase the chance of convergence for CFAs (Kyriazos, 2018), especially for bifactor models (Bader et al, 2022). Alternatively, researchers might continue to explore math anxiety through the subfactors of math evaluation anxiety and learning math anxiety, as this level of analysis may yield significant findings to help explain the discrepancies in the math anxiety research up to this point.

Second, the current study found that a CFA with the theoretically supported 2-factor structure, positive affect and negative affect, did not fit with the data of the current sample. However, post-hoc modifications suggested that this was likely due to the presence of item 8, which did not load well on its proposed factor. Future researchers might explore dropping or modifying item 8, “I find math challenging”, on the MAS-R and examining how this affects the fit of the measure.

Additionally, future researchers might continue to explore the manifestation of math anxiety across secondary and tertiary students, as the AMAS was found to be invariant across school level and thus able to be utilized for both secondary and tertiary students. Further, correlation analysis between variables in this study suggests that age, grade, and school level have a significant relationship with math anxiety.

## Conclusion

Findings of this study indicate that the AMAS is invariant across race and school level, for Black/African American students and White/European American students and secondary and tertiary level students. This suggests that the AMAS can be utilized to compare math anxiety across White/European American and Black/African American students as well as across secondary and tertiary level students.

While results of this study suggest that there are not significant differences in levels of learning math anxiety or math evaluation anxiety across gender or race, math anxiety remains a public issue due to its association with retention in math programs (Ramirez, 2018; Ashcraft & Moore, 2009). Findings from this study suggest that the AMAS can be used to determine levels of math learning anxiety and math evaluation anxiety for Black/African American students. As such, the AMAS can be utilized in secondary and tertiary schools to guide intervention efforts with this population. Various math interventions have been found to be effective with Black/African American tertiary level students, such as bibliotherapy (Herbert & Furner, 1997), mindfulness (Brunyé et al, 2013), and self-affirmations (Samuel & Warner, 2021). The AMAS can be utilized to both inform intervention selection for secondary and tertiary level students as well as determine intervention effectiveness at each of these levels. Further, this study suggests that the AMAS can be used in future research examining differences in learning math anxiety and math evaluation anxiety across Black/African American and White/European American students as well as research focused on the manifestation of math anxiety with Black/African American students.

## REFERENCES


Ahmed, W. (2018). Developmental trajectories of math anxiety during adolescence: associations with STEM career choice. *Journal of Adolescence, 67*(July), 158–166. https://doi.org/10.1016/j.adolescence.2018.06.010

Akin, A., & Kurbanoglu, I. N. (2011). The relationships between math anxiety, math attitudes, and self-efficacy: a structural equation model. *Studia Psychologica*, *3*(53), 263–273.

Allison, P. D. (2012). Statistics and Data Analysis Handling Missing Data by Maximum Likelihood. *SAS Global Forum*, 1–21.

Amato, L. (2018). *Math anxiety: The impact on traditionally underserved and marginalized adult female undergraduate students in elementary statistics*. [Doctoral Dissertation, Northeastern University.] Available from ProQuest Dissertations and Theses database.(UMI No. 10884).

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

Ashcraft, M. H. (2002). Math anxiety: personal, educational, and cognitive consequences. *Current Directions on Psychological Science*, *11*, 181–185. https://doi.org/10.1111/1467-8721.00196

Ashcraft, M. H., & Kirk, E. P. (2001). The relationships among working memory, math anxiety, and performance. *Journal of Experimental Psychology*, *130*(2), 224–237. https://doi.org/10.1037//0096-3445.130.2.224

Ashcraft, M. H., & Moore, A. M. (2009). Mathematics anxiety and the affective drop in performance. *Journal of Psychoeducational Assessment*, *27*(3), 197–205. https://doi.org/10.1177/0734282908330580

Bader, M., Jobst, L. J., & Moshagen, M. (2022). Sample size requirements for bifactor models. *Structural Equation Modeling: A Multidisciplinary Journal*, 1–12. https://doi.org/10.1080/10705511.2021.2019587

Bai, H. (2011). Cross-validating a bidimensional mathematics anxiety scale. *Assessment*, *18*(1), 115–122. https://doi.org/10.1177/1073191110364312

Bauer, D. J. (2005). The Role of Nonlinear Factor-to-Indicator Relationships in Tests of Measurement Equivalence. *American Psychological Association*, *10*(3), 305–316. https://doi.org/10.1037/1082-989X.10.3.305

Beaton, D. E., Bombardier, C., Guillemin, F., & Ferraz, M. B. (2000). Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine*, *25*(24), 3186–3191. https://doi.org/10.1097/00007632-200012150-00014

Benítez, I., Adams, B. G., & He, J. (2019). An integrated approach to bias in a longitudinal survey in the United Kingdom: assessing construct, method, and item bias in the general health questionnaire. *Assessment*, *26*(7), 1194– 1206. https://doi.org/10.1177/1073191117739979

Benítez, I., Vijver, F. Van De, & Padilla, J. (2017). An integrated approach to bias in the Spanish version of the General Health Questionnaire (GHQ-12). *European Journal of Psychological Assessment*, *35*(2), 233–240. https://doi.org/10.1027/1015-5759/a000395

Birgin, O., Balo, M., Çatl, H., & Gürbüz, R. (2010). An investigation of mathematics anxiety among sixth through eighth grade students in Turkey. *Learning and Individual Differences*, *20*, 654–658. https://doi.org/10.1016/j.lindif.2010.04.006

Booker, K. C., & Lim, J. H. (2018). Belongingness and Pedagogy: Engaging African American Girls in Middle School Mathematics. *Youth & Society*, *50*(8), 1037–1055. https://doi.org/10.1177/0044118X16652757

Brunyé, T. T., Mahoney, C. R., Giles, G. E., Rapp, D. N., Taylor, H. A., & Kanarek, R. B. (2013). Learning to relax: evaluating four brief interventions for overcoming the negative emotions accompanying math anxiety. *Learning and Individual Differences*, *27*, 1–7. https://doi.org/10.1016/j.lindif.2013.06.008

Brown, T. A. (2006). Confirmatory factor analysis for applied research. New York, NY: Guilford.

Buratta, L., Piccirilli, M., Lanfaloni, G., Ilicini, S., Bedetti, C., & Elisei, S. (2019). Mathematics anxiety and cognitive performance in adolescent students. *Psychiatria Danubina*, *31*(3), 479–485. PMID: 31488776.

Byrne, B. M., Shavelson, R. J., & Muthén, B. (1989). Testing for the equivalence of factor covariance and mean structures: the issue of partial measurement invariance. *Psychological Bulletin*, *105*(3), 456–466. https://doi.org/10.1037/0033-2909.105.3.456

Capraro, M. M., Capraro, R. M., & Henson, R. K. (2001). Measurement error of scores on the mathematics anxiety rating scale across studies. *Educational and Psychological Measurement*, *61*(3), 373–386. https://doi.org/10.1177/00131640121971266

Carey, E., Hill, F., Devine, A., & Szucs, D. (2017). The modified abbreviated math anxiety scale: A valid and reliable instrument for use with children. *Frontiers in Psychology*, *8*(JAN), 1–13. https://doi.org/10.3389/fpsyg.2017.00011

Carroll, S. (2010). *The Relationship of Math Anxiety and Mathematics Comprehension in Middle School Students*. [Doctoral Dissertation, Walden University] https://doi.org/10.1111/j.1467-8616.2008.00521.x

Caviola, S., Primi, C., Chiesi, F., & Mammarella, I. C. (2017). Psychometric properties of the Abbreviated Math Anxiety Scale (AMAS) in Italian primary school children. *Learning and Individual Differences*, *55*, 174–182. https://doi.org/10.1016/j.lindif.2017.03.006

Cheung, G. W., & Rensvold, R. B. (1998). Cross-cultural comparisons using non-invariant measurement items. *Applied Behavioral Science Review*, 6, 93-110. https://doi.org/10.1016/S1068-8595(99)80006-3

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling*, 14, 464–504. http://dx.doi.org/10.1080/10705510701301834

Chen, F. F. (2008). What happens if we compare chopsticks with forks? The impact of making inappropriate comparisons in cross-cultural research. *Journal of Personality and Social Psychology*, 95, 1005–1018. http://dx.doi.org/10.1037/e514412014-064.

Clark, F. K. (2004). *Sources of Math Anxiety Among failing minority students taking algebra and geometry in high-poverty schools located in SPA 6 of Los Angeles County (California).*[Doctoral Dissertation, Pepperdine University]. Available from ProQuest Dissertations and Theses database.

Dayan, M., Ozer, M., & Almazrouei, H. (2017). The role of functional and demographic diversity on new product creativity and the moderating impact of project uncertainty. *Industrial Marketing Management*, *61*, 144–154. https://doi.org/10.1016/j.indmarman.2016.04.016

Devine, A., Fawcett, K., Szűcs, D., & Dowker, A. (2012). Gender differences in mathematics anxiety and the relation to mathematics performance while controlling for test anxiety. *Behavioral and Brain Functions*, *8*(33), 1–9. https://doi.org/10.1186/1744-9081-8-33

Diedenhofen, B., & Musch, J. (2016). cocron: A Web Interface and R Package for the Statistical Comparison of Cronbach's Alpha Coefficients. *International Journal of Internet Science*, *11*(1).

Dowker, A., Sarkar, A., & Looi, C. Y. (2016). Mathematics anxiety: what have we learned in 60 years? *Frontiers in Psychology*, *7*(April). https://doi.org/10.3389/fpsyg.2016.00508

Dutton, W. H. (1954). Measuring attitudes toward arithmetic. *The Elementary School Journal*, *55*(1), 24–31. https://doi.org/10.1086/458640

Dykeman, C. (2017). The weighted average of Abbreviated Math Anxiety Scale (AMAS) studies on college students.

Else-Quest, N. M., Hyde, J. S., & Linn, M. C. (2010). Cross-national patterns of gender differences in mathematics: a meta-analysis. *Psychological Bulletin*, *136*(1), 103–127. https://doi.org/10.1037/a0018053

Feldt, Leonard S, Woodruff, David J & Salih, Fathi A. (1987). Statistical inference for coefficient alpha. *Applied Psychological Measurement*, 11, 93-103. doi:10.1177/014662168701100107

File, T., & Ryan, C. (2014). Computer and Internet Use in the United States: 2013. In *American Community Survey Reports* (Issue November). http://www.census.gov/prod/2013pubs/p20-569.pdf

Flowerdew, M. R. (2021). *Demographic Predictors of Math Anxiety in Elementary School Children.* [Masters Theses. 216. University of Northern Colorado]. https://digscholarship.unco.edu/theses/216

Fuson, J. Y. (2007). *A study of the relationship between math anxiety and select demographics.* [Doctoral Dissertation, Capella University]. Available from ProQuest Dissertations and Theses database.

Gignac, G. E. (2016). The higher-order model imposes a proportionality constraint: That is why the bifactor model tends to fit better. *Intelligence*, *55*, 57–68. https://doi.org/10.1016/j.intell.2016.01.006

Goodboy, A. K., & Martin, M. M. (2020). Omega over alpha for reliability estimation of unidimensional communication measures. *Annals of the International Communication Association*, *44*(4), 422–439. https://doi.org/10.1080/23808985.2020.1846135

Griggs, M. S., Rimm-Kaufman, S. E., Merritt, E. G., & Patton, C. L. (2013). The responsive classroom approach and fifth grade students' math and science anxiety and self-efficacy. *School Psychology Quarterly*, *28*(4), 360–373. https://doi.org/10.1037/spq0000026

Grother, P., Ngan, M., & Hanaoka, K. (2019). Face Recognition Vendor Test (FRVT) Part 3: Demographic Effects. In *National Institute of Standards and Technology: U.S. Department of Commerce: Vol. December*.

Guilera, G., Gómez-benito, J., Hidalgo, M. D., & Sánchez-meca, J. (2013). Type I error and statistical power of the Mantel-Haenszel procedure for detecting DIF: a meta-analysis. *American Psychological Association*, *18*(4), 553–571. https://doi.org/10.1037/a0034306

Hackett, G. (1985). Role of mathematics self-efficacy in the choice of math-related majors of college women and men: a path analysis. *Journal of Counseling Psychology*, *32*(1), 47–56. https://doi.org/10.1037/0022-0167.32.1.47

Hayes, A. F., & Coutts, J. J. (2020). Use Omega rather than Cronbach's Alpha for estimating reliability. *Communication Methods and Measures*, *14*(1), 1–24. https://doi.org/10.1080/19312458.2020.1718629

He, H. (2007). *Adolescents' Perception of Parental and Peer Mathematics Anxiety and Attitude Toward Mathematics: A Comparative Study of European-American and Mainland-Chinese Students*. [Doctoral Dissertation, Washington State University]. Available from ProQuest Dissertations and Theses database.

Hebert, T. I., & Furner, J. M. (1997). Helping high ability students overcome math anxiety through Bibliotherapy. *The Journal of Secondary Gifted Education*, *VIII*(4), 164–178. https://doi.org/10.1177/1932202X9700800403

Hembree, R. (1990). The nature, effects, and relief of mathematics anxiety. *Journal for Research in Mathematics Education*, *21*(1), 33–46. https://doi.org/10.2307/749455

Hendel, D. D., & Davis, S. O. (1978). Effectiveness of an intervention strategy for reducing mathematics anxiety. *Journal of Counseling Psychology*, *25*(5), 429–434. https://doi.org/10.1037/0022-0167.25.5.429

Hill, F., Mammarella, I. C., Devine, A., Caviola, S., Passolunghi, M. C., & Szucs, D. (2016). Maths anxiety in primary and secondary school students: Gender differences, developmental changes and anxiety specificity. *Learning and Individual Differences*, *48*, 45–53. https://doi.org/10.1016/j.lindif.2016.02.006

Honey, M., Alberts, B., Bass, H., Castillo, C., Lee, O., Strutchens, M. M., Vermillion, L., & Rodriguez, F. (2020). *STEM Education for the Future: A Visioning Report* (Vol. 1, Issue 2018).

Hong, L., & Page, S. E. (2004). Groups of diverse problem solvers can outperform groups of high-ability problem solvers. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(46), 16385–16389. https://doi.org/10.1073/pnas.0403723101

Hopko, D. R. (2003). Confirmatory factor analysis of the math anxiety rating scale-revised. *Educational and Psychological Measurement*, *63*(2), 336–351. https://doi.org/10.1177/0013164402251041

Hopko, D. R., Bare, R. L., & Hunt, M. K. (2003). The Abbreviated Math Anxiety Scale (AMAS) Construction, Validity, and Reliability. *Assessment*, *10*(2), 178–182. https://doi.org/10.1177/1073191103252351

Horrigan, J. (2009). Wireless Internet. In *Pew Internet Research* (Issue July). https://doi.org/10.1201/noe1420043266.ch149

Hoyle, R. H., & Gottfredson, N. C. (2015). Sample size considerations in prevention research applications of multilevel modeling and structural equation modeling. *Prevention Science*, *16*(7), 987–996. https://doi.org/10.1007/s11121-014-0489-8

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Huang, X., Zhang, J., & Hudson, L. (2019). Impact of math self-efficacy, math anxiety, and growth mindset on math and science career interest for middle school students: the gender moderating effect. *European Journal of Psychology of Education*, *34*(3), 621–640. https://doi.org/10.1007/s10212-018-0403-z

Iammartino, R., Bischoff, J., Willy, C., & Shapiro, P. (2016). Emergence in the U.S. Science, Technology, Engineering, and Mathematics (STEM) workforce: an agent-based model of worker attrition and group size in high-density STEM organizations. *Complex & Intelligent Systems*, *2*(1), 23–34. https://doi.org/10.1007/s40747-016-0015-7

Ibrahim, H., Johnson, O. School Discipline, Race–Gender and STEM Readiness: A Hierarchical Analysis of the Impact of School Discipline on Math Achievement in High School. *The Urban Review* 52, 75–99 (2020). https://doi.org/10.1007/s11256-019-00513-6

Iossi, L. (2007). Strategies for reducing math anxiety in post-secondary students. In S. M. Nielsen & M. S. Plakhotnik (Eds.), *Proceedings of the Sixth Annual College of Education Research Conference: Urban and International Education Section* (pp. 30-35). Miami: Florida International University. http://coeweb.fiu.edu/research_conference/

Iossi, L. H. (2009). *The mathematics anxiety of bilingual community college students* [Doctoral Dissertation, Florida International University]. https://doi.org/10.25148/etd.FI10022525

Jakobsen, J.C., Gluud, C., Wetterslev, J. et al. When and how should multiple imputation be used for handling missing data in randomised clinical trials – a practical guide with flowcharts. *BMC Med Res Methodol 17*, 162 (2017). https://doi.org/10.1186/s12874-017-0442-1

Jazdzewski, K. (2011). *Attitudes about Mathematics: Compare and Contrast Boys and Girls from High and Low Socio-economic Status*. Masters thesis, California State University

Johnson, D. R., & Young, R. (2011). Toward best practices in analyzing datasets with missing data: comparisons and recommendations. *Journal of Marriage and Family*, *73*(5), 926–945. https://doi.org/10.1111/j.1741-3737.2011.00861.x

Jöreskog, Karl G. (1971). Simultaneous Factor Analysis in Several Populations. *Psychometrika* 36:409–426. https://doi.org/10.1007/BF02291366

Kerr, B., & Kurpius, S. E. R. (2004). Encouraging talented girls in math and science: effects of a guidance intervention. *High Ability Studies*, *15*(1), 85–102. https://doi.org/10.1080/1359813042000225357

Kim, H. Y. (2013). Statistical notes for clinical researchers: assessing normal distribution (2) using skewness and kurtosis. Restorative Dentistry & Endodontics, 38(1), 52-54. doi:10.5395/rde.2013.38.1.52

Kyriazos, T. A. (2018). Applied psychometrics: sample size and sample power considerations in factor analysis (EFA, CFA) and SEM in general. *Psychology*, *9*, 2207–2230. https://doi.org/10.4236/psych.2018.98126

Kyttälä, M., & Björn, P. M. (2014). The role of literacy skills in adolescents' mathematics word problem performance: controlling for visuo-spatial ability and mathematics anxiety. *Learning and Individual Differences*, *29*, 59–66. https://doi.org/10.1016/j.lindif.2013.10.010

LaGue, A., Eakin, G., & Dykeman, C. (2019). The impact of mindfulness-based cognitive therapy on math anxiety in adolescents. *Preventing School Failure, 63*(2), 142–148. https://doi.org/10.1080/1045988X.2018.1528966

Lang, K.M., Little, T.D. Principled Missing Data Treatments. *Prevention Science*, *19*, 284–294 (2018). https://doi.org/10.1007/s11121-016-0644-5

Lee, J. (2009). Universals and specifics of math self-concept, math self-efficacy, and math anxiety across 41 PISA 2003 participating countries. *Learning and Individual*

*Differences*, *19*(3), 355–365. https://doi.org/10.1016/j.lindif.2008.10.009

Liang, X., & Yang, Y. (2014). An evaluation of WLSMV and Bayesian methods for confirmatory factor analysis with categorical indicators. *International Journal of Quantitative Research in Education*, *2*(1), 17. https://doi.org/10.1504/ijqre.2014.060972

Luttenberger, S., Wimmer, S., & Paechter, M. (2018). Spotlight on math anxiety. *Psychology Research and Behavior Management*, 311–322. https://doi.org/10.2147/PRBM.S141421

Major, D. A., Holland, J. M., & Oborn, K. L. (2012). The influence of proactive personality and coping on commitment to STEM majors. *Career Development Quarterly*, *60*(1), 16–24. https://doi.org/10.1002/j.2161-0045.2012.00002.x

Maloney, E. A., Risko, E. F., Ansari, D., & Fugelsang, J. (2010). Mathematics anxiety affects counting but not subitizing during visual enumeration. *Cognition, 114*, 293-297. http://dx.doi.org/10.1016/j.cognition.2009.09.013

Mau, W. C. J., & Li, J. (2018). Factors influencing STEM career aspirations of underrepresented high school students. *Career Development Quarterly*, *66*(3), 246–258. https://doi.org/10.1002/cdq.12146

McCormick, M. (1998). *The influence of gender-role identity, mathematics self-efficacy, an outcome expectations on the math- and science-related career interests of gifted adolescent girls.* (Publication No. 9706100) [Doctoral Dissertation, University of Utah]. ProQuest Dissertations Publishing.

Meade, A. W., Johnson, E. C., & Braddy, P. W. (2008). Power and sensitivity of alternative fit indices in tests of measurement invariance. *Journal of Applied Psychology*, *93*(3), 568–592. https://doi.org/10.1037/0021-9010.93.3.568

Meredith, W. (1964). Rotation to achieve factorial invariance. *Psyctwmetrika, 19,* 187-206. Retrieved June 25, 2021, from https://link.springer.com/content/pdf/10.1007/BF02289700.pdf

Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika*, *58*(4), 525-543. https://doi-org.libproxy.temple.edu/10.1007/BF02294825

Meredith, W., & Teresi, J. A. (2006). An essay on measurement and factorial invariance. *Medical care*, S69-S77. Retrieved June 25, 2021, from http://www.jstor.org/stable/41219507

Merritt, W. (2011). *Exploring math anxiety as it relates to math achievement, gender, and race.* (Publication No. 3487164) [Doctoral Dissertation, Mississippi State University]. ProQuest Dissertations Publishing.

Morsanyi, K., Busdraghi, C., & Primi, C. (2014). Mathematical anxiety is linked to reduced cognitive reflection: a potential road from discomfort in the mathematics classroom to susceptibility to biases. *Behavioral and Brain Functions*, *10*. https://doi.org/10.1186/1744-9081-10-31

Mulaik, S. A. (2009). *Linear causal modeling with structural equations.* CRC Press.

National Science Foundation, N. C. for S. and E. S. (2019). *Women, Minorities, and Persons with Disabilities in Science and Engineering: 2019*. Special Report NSF 19-304. Alexandria, VA. https://www.nsf.gov/statistics/wmpd

Østergaard, C. R., Timmermans, B., & Kristinsson, K. (2011). Does a different view create something new? The effect of employee diversity on innovation. *Research Policy*, *40*(3), 500–509. https://doi.org/10.1016/j.respol.2010.11.004

Park, D., Ramirez, G., & Beilock, S. L. (2014). The role of expressive writing in math anxiety. *Journal of Experimental Psychology: Applied*, *20*(2), 103–111. https://doi.org/10.1037/xap0000013

Pendergast, L. L., Embse, N. Von Der, Kilgus, S. P., & Eklund, K. R. (2017). Measurement equivalence: a non-technical primer on categorical multi-group confirmatory factor analysis in school psychology. *Journal of School Psychology*, *60*, 65–82. https://doi.org/10.1016/j.jsp.2016.11.002

Perry, A. B. (2004). Decreasing math anxiety in college students. *College Student Journal*.

Plake, B. S., & Parker, C. S. (1982). The development and validation of a revised version of the mathematics anxiety rating scale. *Educational and Psychological Measurement*, *42*. https://doi.org/10.1177/001316448204200218

Pletzer, B., Kronbichler, M., & Nuerk, H. (2015). Mathematics anxiety reduces default mode network deactivation in response to numerical tasks. *Frontiers in Human Neuroscience*, *9*(April), 1–12. https://doi.org/10.3389/fnhum.2015.00202

Pletzer, B., Wood, G., Scherndl, T., Kerschbaum, H. H., & Nuerk, H. C. (2016). Components of mathematics anxiety: factor modeling of the MARS30-brief. *Frontiers in Psychology*, *7*(FEB), 1–14. https://doi.org/10.3389/fpsyg.2016.00091

Primi, C., Busdraghi, C., Tomasetto, C., Morsanyi, K., & Chiesi, F. (2014). Measuring math anxiety in Italian college and high school students: validity, reliability and gender invariance of the Abbreviated Math Anxiety Scale (AMAS). *Learning and Individual Differences*, *34*, 51–56. https://doi.org/10.1016/j.lindif.2014.05.012

Punaro, L., & Reeve, R. (2012). Relationships between 9-Year-Olds' math and literacy worries and academic abilities. *Child Development Research*, *2012*, 1–11. https://doi.org/10.1155/2012/359089

Ramirez, G., Gunderson, E. A., Levine, S. C., & Beilock, S. L. (2013). Math anxiety, working memory, and math achievement in early elementary school. *Journal of Cognition and Development*, *14*(2), 187–202. https://doi.org/10.1080/15248372.2012.664593

Ramirez, G., Shaw, S. T., Maloney, E. A., Ramirez, G., Shaw, S. T., Maloney, E. A., Anxiety, M., Ramirez, G., Shaw, S. T., & Maloney, E. A. (2018). Math anxiety: past research, promising interventions, and a new interpretation framework. *Educational Psychologist*, *53*(3), 145–164. https://doi.org/10.1080/00461520.2018.1447384

Reise, S. P., Widaman, K. F., & Pugh, R. H. (1993). Confirmatory factor-analysis and item response theory: Two approaches for exploring measurement invariance. *Psychological Bulletin, 114*, 552-566.

Richardson, F. C., & Suinn, R. M. (1972). The mathematics anxiety rating scale: psychometric data. *Journal of Counseling Psychology*, *18*(6), 6–9. http://dx.doi.org/10.1037/h0033456

Samuel, T. S., & Warner, J. (2021). "I Can Math!": Reducing math anxiety and increasing math self-efficacy using a mindfulness and growth mindset-based intervention in first-year students. *Community College Journal of Research and Practice*, *45*(3), 205–222. https://doi.org/10.1080/10668926.2019.1666063

Sarkar, A., Dowker, A., & Kadosh, R. C. (2014). Cognitive enhancement or cognitive cost: trait-specific outcomes of brain stimulation in the case of mathematics anxiety. *The Journal of Neuroscience*, *34*(50), 16605–16610. https://doi.org/10.1523/JNEUROSCI.3129-14.2014

Shi, D., Song, H., & Lewis, M. D. (2019). The impact of partial factorial invariance on cross-group comparisons. *Assessment*, *26*(7), 1217–1233. https://doi.org/10.1177/1073191117711020

Steele, C. (2010). *Whistling Vivaldi and other clues to how stereotypes affect us* (1st ed.). W. W. Norton & Company, Inc.

Steenkamp, J. B. E. M., & Baumgartner, H. (1998). Assessing measurement invariance in cross-national consumer research. *Journal of Consumer Research*, *25*(1), 78–90. https://doi.org/10.1086/209528

Suinn, R. M., & Edwards, R. (1982). The measurement of mathematics anxiety: the mathematics anxiety rating scale for adolescents - MARS-A. *Journal of Clinical Psychology*, *88*(8), 576–581. https://doi.org/10.1002/1097-4679(198207)38:3<576::AID-JCLP2270380317>3.0.CO;2-V

Tapia, M., & Marsh, G. E. (2000). Attitudes toward mathematics instrument: an investigation with middle school students. *Annual Meeting of the Mid-South Educational Research Association*, *1*, 1–32. papers3://publication/uuid/A0A266F4-18E4-41C4-B1F6-EDFB58C6E8CE

Tobias, S. (1990). Math anxiety: an update. *NACADA Journal*, *10*(1). https://doi.org/10.12930/0271-9517-10.1.47

Toland, M. D., & Usher, E. L. (2016). Assessing mathematics self-efficacy: how many categories do we really need? *Journal of Early Adolescence*, *36*(7), 932–960. https://doi.org/10.1177/0272431615588952

Van Buuren, S., & Groothuis-Oudshoorn, K. (2011). Mice: Multivariate Imputation by Chained Equations in R. *Journal of Statistical Software*, 45(3), 1-67. DOI 10.18637/jss.v045.i03.

Van Buuren, S., Brand, J. P.L., Groothuis-Oudshoorn,C. B. M., & Rubin, D. B. (2006) Fully conditional specification in multivariate imputation, *Journal of Statistical*

*Computation and Simulation*, 76:12, 1049-1064, DOI: 10.1080/10629360600810434

Varelas, M., D. B., Martin, and J. M. Kane. 2013. Content learning and identity construction: A framework to strengthen African American students' mathematics and science learning in urban elementary schools. *Human Development* 55 (5–6): 319–339

Vijver, V. De, & Tanzer, N. K. (2004). Bias and equivalence in cross-cultural assessment: An overview. *European Review of Applied Psychology*, *54*, 119–135. https://doi.org/10.1016/j.erap.2003.12.004

Wang, M., Deng, W., Hu, J., Tao, X., & Huang, Y. (2019). Racial faces in the wild: Reducing racial bias by information maximization adaptation network. *Proceedings of the IEEE International Conference on Computer Vision*, 692–702. https://doi.org/10.1109/ICCV.2019.00078

Wang, Z., Rimfeld, K., Shakeshaft, N., Schofield, K., & Malanchini, M. (2020). The longitudinal role of mathematics anxiety in mathematics development: Issues of gender differences and domain-specificity. *Journal of Adolescence*, *80*(March), 220–232. https://doi.org/10.1016/j.adolescence.2020.03.003

Xie, F., Xin, Z., Chen, X., & Zhang, L. (2019). Gender Difference of Chinese High School Students' Math Anxiety: The Effects of Self-Esteem, Test Anxiety and General Anxiety. *Sex Roles*, *81*, 235–244. https://doi.org/10.1007/s11199-018-0982-9

Xue, Y., & Larson, R. (2015). STEM crisis or STEM surplus? Yes and yes. In *Monthly Labor Review*. U.S. Bureau of Labor Statistics, May 2015, https://doi.org/10.21916/mlr.2015.14.

Yardi, S., & Bruckman, A. (2012). Income, race, and class: Exploring socioeconomic differences in family technology use. *Conference on Human Factors in Computing Systems - Proceedings*, 3041–3050. https://doi.org/10.1145/2207676.2208716

Young, C. B., Wu, S. S., & Menon, V. (2012). The Neurodevelopmental Basis of Math Anxiety. *Psychological Science*, *23*(5), 492–501. https://doi.org/10.1177/0956797611429134.The

Young, J. R., & Young, J. L. (2015). Anxious for answers: A meta-analysis of the effects of anxiety on African American K-12 students' mathematics achievement. *Journal of Mathematics Education at Teacher's College*, *6*(2). https://doi.org/10.7916/jmetc.v6i2.611

Young, J. R., & Young, J. L. (2016). Young, Black, and anxious: describing the Black student mathematics anxiety research using confidence intervals. *Journal of Urban Mathematics Education*, *9*(1), 79–93. http://education.gsu.edu/JUME

Young, J. L, Young, J. R., & Capraro, M. M. (2017) Black Girls' Achievement in Middle Grades Mathematics: How Can Socializing Agents Help?, *The Clearing House: A Journal of Educational Strategies, Issues and Ideas*, 90:3, 70-76, DOI: 10.1080/00098655.2016.1270657

Yüksel-Şahin, F. (2008). Mathematics anxiety among 4th and 5th grade Turkish elementary school students. *International Electronic Journal of Mathematics Education*, *3*(3), 179–192. https://doi.org/10.1016/j.sbspro.2012.06.362

Zeldin, A. L., & Pajares, F. (2016). Against the odds: self-efficacy beliefs of women in mathematical, scientific, and technological careers. *American Educational Research Journal*, *37*(1), 215–246. https://doi.org/10.2307/1163477

## APPENDIX A

DIF PREDICTIONS

**Table 12. DIF Predictions**

| Measure | Item | DIF |
|---|---|---|
| AMAS | 3. Watching a teacher work an algebraic equation on the blackboard | Previous research utilizing the MARS-A, the predecessor of the AMAS, suggested that African American students rated themselves more highly than a control group on items related to classroom culture and procedures (Clark, 2004). This suggests that the classroom culture and procedures may be the cause of the anxiety these students are recognizing and reporting on this measure. African American students then might rate themselves more highly on this item than European American students due to differences in experiences of classroom culture and procedures as opposed to true differences in math anxiety. Additionally, qualitative research using a phenomenological approach with semi-structured interviews (Belenky et. al. 1986) suggests that gender differences in learning process and perceptions of the classroom climate contributes to girls' alienation and lack of self-confidence in the classroom. Further, quantitative data (American Association of University Women, 1992) suggests that girls receive significantly less attention from classroom teachers than boys, African American girls have fewer interactions with teachers than do white girls (despite evidence that they attempt to initiate interactions more frequently), and sexual harassment of girls by boys in schools in the U.S. is increasing. This poor school climate may lead girls to rate themselves more highly on item 3 on the AMAS, as it pertains to the classroom environment, instead of rating this item more highly due to their experience of math anxiety. |

Table 12. (continued)

| Measure | Item | DIF |
|---|---|---|
| AMAS | 6. Listening to a lecture in math class | Similar to the previous item, item 6 may be rated more highly among African American students because the classroom culture and procedures may be more anxiety provoking among African American students (Clark 2004), rather than due to their experiences of math anxiety.<br><br>Additionally, this item may be rated more highly among girls due to their perceptions of classroom climate and learning process, rather than due to their experience of math anxiety (Belenky et. al. 1986; American Association of University Women, 1992). |
| MAS-R | 3. I think that I will use math in the future | Item 3 on the MAS-R may not be rated as highly by African American students and girls due to lack of representation of African Americans and women in STEM fields (Major, 2012; National Science Foundation, 2019). Representation in STEM fields is a predictor of marginalized students' choice to pursue a STEM field (Kerr and Kurpius, 2004). African Americans and girls may rate themselves lower on this item than European Americans and boys because they may be less likely to see themselves as using math in their careers in the future, instead of because of their experiences of math anxiety or negative affect towards math. |
| MAS-R | 5. Math relates to my life | Similar to item 3 on the MAS-R, item 5 may not be rated as highly by African Americans and girls as European Americans and boys because they may not see themselves as using math in their careers due to a lack of representation of African Americans and women in STEM (Major, 2012; National Science Foundation, 2019) and the positive association between representation and STEM career choice in marginalized students (Kerr & Kurpius, 2004), instead of due to the experience of math anxiety or negative affect towards math. |

## APPENDIX B

POST-HOC MODIFICATION PREDICTIONS

**Table 13. Post-hoc Modification Predictions**

| Measure | Item/Factor | Modification |
|---|---|---|
| AMAS | Item 5. Being given a homework assignment of many difficult problems that is due the next class meeting | This item theoretically, and in past research (Hopko, 2003), loads onto the Math Evaluation Anxiety (MEA) factor on the AMAS. This factor was originally obtained by Plake and Parker (1982) through a principal component analysis (PCA) of their 24-item revision (R-MARS) of the theoretically unidimensional MARS (Richardson & Suinn, 1972). A post-hoc interpretation of the MEA factor of math anxiety on the R-MARS concludes that MEA is associated with the evaluation of one's math ability (Plake & Parker, 1982).<br><br>The other three items on the MEA factor of the AMAS are specific to math tests and quizzes: (2) Thinking about a math test that will be given the following day, (4) Taking an evaluation in a math class, and (8) Being given a pop quiz in a math class. Item 5, however, asks the rater if they would be made anxious by being given a homework assignment. As item 5 is conceptually different from Items 2, 4, and 8, in that it does not discuss a test/quiz, and the MEA factor was developed through post-hoc interpretation of a PCA, it may lead to poor model fit for the AMAS when compared to the data gathered in this study. |

Table 13. (continued)

| Measure | Item/Factor | Modification |
|---|---|---|
| AMAS | Item 5. Being given a homework assignment of many difficult problems that is due the next class meeting | Notably, at the conception of the AMAS, Hopko (2003) obtained a factor loading of .66 for item 5 on the MEA factor, while factor loadings for the other three items of MEA were .86, .89, and .84.<br><br>Given the difference in conceptualization of item 5 with items 2, 4, and 8 as well as their difference in factor loadings on the MEA factor, if modification indices suggest post-hoc modifications for parameters related to item 5, such as error correlations between item 5 and other items on the AMAS, and item 5's factor loadings onto MEA and Math Anxiety, such modifications will be pursued. |
| AMAS | Learning Math Anxiety | The learning math anxiety (LMA) factor on the AMAS is composed of items 1) Having to use the tables in the back of a math book, 3) Watching a teacher work an algebraic equation on the blackboard, 6) Listening to a lecture in a math class, 7) Listening to another student explain a math formula, and 9) Starting a new chapter in a math book. The LMA factor was originally obtained by Plake and Parker (1982) through a principal component analysis (PCA) of their 24-item revision (the R-MARS) of the theoretically unidimensional MARS (Richardson & Suinn, 1972). A post-hoc interpretation of the LMA factor of math anxiety on the R-MARS concludes that LMA is associated with the activity or process of studying statistics (Plake & Parker, 1982).<br><br>While items 3, 6, and 7 ask the rater if they would be made anxious by a description of learning math from a teacher or another student, items 1 and 9 ask the rater if they would be made anxious by a description of |

Table 13. (continued)

| Measure | Item/Factor | Modification |
|---|---|---|
| AMAS | Learning Math Anxiety | using a math textbook. Conceptually, these are two different ways of learning. Due to the theoretical difference of the types of learning that are presented in the items on the LMA factor, the LMA factor may lead to poor fit of the model to the data in this study.<br><br>Notably, when creating the MARS, the predecessor of the R-MARS and AMAS, Richardson and Suinn (1972) conceptualized math anxiety as a unidimensional construct. Although their data suggested two components of math anxiety, Plake and Parker (1982) did not have a theoretical basis for the division of their measure into two separate factors (MEA and LMA).<br><br>As the LMA factor has no theoretical basis, and its items appear to be theoretically split between two separate ways of learning, if modification indices suggest post-hoc modifications that would alter the LMA factor, for example splitting it into two factors that represent learning in the classroom and from a textbook, those modifications will be pursued. |
| MAS-R | Item 3. I think that I will use math in the future | Item 3 has been found through PCA to fall within the positive affect factor of the MAS-R (Bai, 2009). However, the strength of this loading has failed to replicate (Bai, 2011). This item had the lowest factor loading (.39) on the positive affect subscale of the MAS-R on a principal components analysis conducted on the measure in past research (Bai, 2011). Factor loadings from the other items on the positive affect subscale in Bai (2011)'s PCA ranged from .55 to .85. This suggests that Item 3 is less related to the other items. |

Table 13 (continued)

| Measure | Item/Factor | Modification |
|---|---|---|
| MAS-R | Item 3. I think that I will use math in the future | on the positive affect subscale of this measure.<br><br>Conceptually, this item is different from other items on the measure that ask the rater about their enjoyment and interest in math: 13) I enjoy learning with mathematics, 1) I find math interesting, 12) math is one of my favorite subjects.<br><br>As Item 3 is conceptually different from other items on the positive affect factor, and the strength of its loading has not replicated in PCA (Bai, 2011), if modification indices suggest that freeing parameters related to this item would increase the fit of the MAS-R model to the data, such post-hoc modifications will be pursued. These parameters might include correlations between item 3 and other items on the positive affect factor, item 3's factor loading onto the positive affect factor and the math anxiety factor on the MAS-R, or the addition of a third factor that includes item 3 on the MAS-R |
| MAS-R | Item 5. Math relates to my life | Similar to item 3 on the positive affect factor of the MAS-R, item 5 is conceptually different from items 1, 12, and 13. While items 1, 12, and 13 ask raters about their enjoyment and interest in math, item 5 asks raters the extent to which math relates to their lives. A PCA has extracted a factor loading of .55 for item 5 onto the positive affect factor of the MAS-R (Bai, 2011). Items 1, 12, and 13 have factor loadings of .83, .79, and .85 on the positive affect factor, respectively. Items 3 and 5 may be more associated with each other than the other items in the positive affect factor because they are both asking raters about the relevance of math outside |

Table 13 (continued)

| Measure | Item/Factor | Modification |
|---|---|---|
| MAS-R | Item 5. Math relates to my life | of the school system, as opposed to asking about enjoyment and interest in math.<br><br>Due to this conceptual difference, if modification indices in this study suggest that freeing parameters related to item 5 would increase the fit of the MAS-R model to the data, such post-hoc modifications will be pursued. These parameters might include correlations between item 5 and other items on the positive affect factor, factor loadings on positive affect and math anxiety on this measure, and/or the addition of a third factor comprised of items 3 and 5. |

## APPENDIX C

ADDITIONAL HIERARCHICAL REGRESSION TABLES

**Table 14.** Hierarchical Linear Regression Model 1 with Gender and Race as Predictors of Learning Math Anxiety

| Coefficient | Estimate | SE | p-value |
|---|---|---|---|
| Intercept | 14.2975 | 0.7298 | <2e-16** |
| Gender | 0.1687 | 0.3616 | 0.641 |
| Race | 0.3454 | 0.3748 | 0.357 |

*Note:* $F(2,636) = 0.662$, $p = 0.5162$, $R^2 = 0.002078$, $R^2_{adj} = -0.001061$, * $p < .05$, ** $p < .01$

**Table 15.** Hierarchical Linear Regression Model 2 with Gender, Race, FORPL, Mother Education Level, and Father Education Level as Predictors of Learning Math Anxiety

| Coefficient | Estimate | SE | p-value |
|---|---|---|---|
| Intercept | 14.72093 | 0.95449 | <2e-16** |
| Gender | 0.11221 | 0.35947 | 0.75502 |
| Race | 0.17986 | 0.38082 | 0.63688 |
| FORPL | -1.20292 | 0.39014 | 0.00214 ** |
| Mother Education Level | 0.38733 | 0.21104 | 0.06692 |
| Father Education Level | 0.03333 | 0.21665 | 0.87779 |

*Note:* $F(5,633) = 3.885$, $p = 0.001781$, $R^2 = 0.02977$, $R^2_{adj} = 0.02211$, * $p < .05$, ** $p < .01$

**Table 16.** Hierarchical Linear Regression Model 1 with Gender and Race as Predictors of Math Evaluation Anxiety

| Coefficient | Estimate | SE | p-value |
|---|---|---|---|
| Intercept | 12.4276 | 0.5625 | <2e-16** |
| Gender | 0.1504 | 0.2787 | 0.590 |
| Race | -0.2305 | 0.2889 | 0.425 |

*Note:* $F(2,636) = 0.3873$, $p = 0.679$, $R^2 = 0.001217$, $R^2_{adj} = -0.001924$, * $p < .05$, ** $p < .01$

**Table 17.** Hierarchical Linear Regression Model 2 with Gender, Race, FORPL, Mother Education Level, and Father Education Level as Predictors of Math Evaluation Anxiety

| Coefficient | Estimate | SE | p-value |
|---|---|---|---|
| Intercept | 12.09604 | 0.74445 | <2e-16** |
| Gender | 0.10351 | 0.28037 | 0.712 |
| Race | -0.28035 | 0.29702 | 0.346 |
| FORPL | -0.03259 | 0.30429 | 0.915 |
| Mother Education Level | 0.24944 | 0.16460 | 0.130 |
| Father Education Level | -0.09368 | 0.16898 | 0.579 |

*Note:* $F(5,633) = 0.7274$, $p = 0.6031$, $R^2 = 0.005713$, $R^2_{adj} = -0.002141$, * $p < .05$, ** $p < .01$

APPENDIX D

INTERCEPTS OF AMAS ITEMS BY SCHOOL LEVEL

**Table 18.** Intercepts of AMAS Items by School Level

| Item | Secondary Group Intercept | Tertiary Group Intercept |
|---|---|---|
| AMAS1 | 2.572 | 2.791 |
| AMAS2 | 2.844 | 2.939 |
| AMAS3 | 2.770 | 2.778 |
| AMAS4 | 3.082 | 2.888 |
| AMAS5 | 2.900 | 2.837 |
| AMAS6 | 2.775 | 2.559 |
| AMAS7 | 2.736 | 2.525 |
| AMAS8 | 2.846 | 2.750 |
| AMAS9 | 3.020 | 2.728 |

## APPENDIX E

MEAN SCORES BY RACE AND GENDER FOR AMAS AND MASR SCALES

**Table 19.** Mean Scores by Race and Gender for AMAS and MASR scales.

| Scale | Sample | Mean |
|---|---|---|
| AMAS | Total | 27.46 |
| AMAS | Black/African American | 27.24 |
| AMAS | White/European American | 27.51 |
| AMAS | Asian American | 27.11 |
| AMAS | Native American | 28.39 |
| AMAS | Male | 27.16 |
| AMAS | Female | 27.82 |
| MASR | Total | 41.00 |
| MASR | Black/African American | 41.17 |
| MASR | White/European American | 40.96 |
| MASR | Asian American | 40.96 |
| MASR | Native American | 39.75 |
| MASR | Male | 40.82 |
| MASR | Female | 41.26 |