Summer 1-1-2016

# Differences in Peer Bullying Victimization by Race and Gender: The Impact of Proficiency, Enjoyment, and Confidence in Math and Science

Daniel S. Wells

# DIFFERENCES IN PEER BULLYING VICTIMIZATION BY RACE AND GENDER: THE IMPACT OF PROFICIENCY, ENJOYMENT, AND CONFIDENCE IN MATH AND SCIENCE

A Dissertation

Submitted to the School of Education

Duquesne University

In partial fulfillment of the requirements for

the degree of Doctor of Philosophy

By

Daniel S. Wells, M.S.Ed.

August 2016

**DUQUESNE UNIVERSITY**
**School of Education**
**Department of Counseling, Psychology, and Special Education**

***Dissertation***

Submitted in partial fulfillment of the requirements for the degree
Doctor of Philosophy (Ph.D.)

**School Psychology Doctoral Program**

***Presented by:***

Daniel S. Wells
B.S. Psychology, Grove City College, 2011
M.S.Ed. Child Psychology, Duquesne University, 2012

**June 22, 2016**

DIFFERENCES IN PEER BULLYING VICTIMIZATION BY RACE AND GENDER: THE IMPACT OF PROFICIENCY, ENJOYMENT, AND CONFIDENCE IN MATH AND SCIENCE

***Approved by:***

_____________________________________________, Chair
Laura M. Crothers, D.Ed.
Professor
Department of Counseling, Psychology, and Special Education
Duquesne University

_____________________________________________, Member
Ara J. Schmitt, Ph.D.
Associate Professor
Department of Counseling, Psychology, and Special Education
Duquesne University

_____________________________________________, Member
James B. Schreiber, Ph.D.
Professor
Department of Educational Foundations and Leadership
Duquesne University

_____________________________________________, Member
Jered Kolbert, Ph.D.
Professor
Department of Counseling, Psychology, and Special Education
Duquesne University

# ABSTRACT

DIFFERENCES IN PEER BULLYING VICTIMIZATION BY RACE AND GENDER: THE IMPACT OF PROFICIENCY, ENJOYMENT, AND CONFIDENCE IN MATH AND SCIENCE

By

Daniel S. Wells

August 2016

Dissertation supervised by Dr. Laura M. Crothers

Despite a significant amount of study and intervention, racial minority students continue to perform at a lower level than their White peers, while female students lag behind their male peers, in terms of math and science achievement. The consistency and resiliency of this achievement gap suggests that these patterns of performance may have become societal expectations. As minority and female students attempt to increase their level of math and science academic performance, and, therefore, violate societal expectations, they may experience a higher risk of another pervasive problem: peer bullying victimization. Previous research has demonstrated that academic success, stereotype violation, race, and gender have all been associated with the experience of bullying. Using the 2011 $8^{th}$ grade, United States sample of the Trends in International Math and Science Study (TIMSS 2011), the current study attempts to determine if racial minority students report higher rates of bullying victimization than White students, and if female students report higher rates of bullying victimization than male students, after controlling for the effects of math and science ability, enjoyment, and confidence. Results indicate that female students do report significantly higher rates of victimization than their male peers, while Hispanic students report significantly higher rates of victimization than their White and Multi-Racial peers. However, while significant differences were shown to exist, those differences represented trivial effect sizes and, therefore, appear to have little noticeable impact on students' bullying experiences. Results related to supplementary analyses, as well as limitations and implications for future research are also discussed.

## DEDICATION

First and foremost, I want to dedicate this dissertation to my wife, Sarah. Sarah, I could not have made it through graduate school and this dissertation process without your steadfast love and support. You were a constant companion, encourager, and positive example throughout this entire process. I will be forever grateful for your dedication to me and to your own academic work, even when you wanted to be doing anything else. With this complete, I am so excited to see all the other things we are going to do together.

I also want to dedicate this to my family, Steve, Lynn, and Abby Wells. Mom and Dad, I am amazed by how much my choice of profession, outlook on life, and skills are clearly a combination of the two of you. I feel so blessed by the laughter of our house, the time both of you spent cultivating my interests and skills, and your daily, unconditional love. Abby Joy, I am so inspired by your selfless love for others and your constant hard work that makes others' lives, including my own, better in so many practical ways. I could not have asked for a more kind, giving, and knuckle-headed sister. I continue to love every minute of growing up with you.

# ACKNOWLEDGEMENT

This dissertation, and my graduate school experience in general, would not have been possible without the consistent support of Dr. Laura Crothers. Dr. Crothers, looking back at my time at Duquesne, your willingness to provide me with opportunities to stretch my own research and academic skills, while still making yourself available for guidance and words of encouragement when needed, significantly impacted my abilities as well as my confidence in those abilities. You served as an excellent role model and mentor to me, and I hope to follow your example of becoming an impactful researcher as well as someone who is invested in the development and success of others.

I am also forever indebted to my dissertation committee members. Dr. Kolbert, thank you for adding your expertise in bullying to my project and for your feedback throughout the writing process. Dr. Schreiber, I truly appreciate your guidance and instruction related to statistics. Your introduction to higher-level statistical procedures will undoubtedly serve me well throughout my career. Finally, Dr. Schmitt, your flexibility, expertise in so many areas of school psychology, and willingness to help me throughout this process was essential to me reaching this point in my research and applied practice. Thank you all so much.

TABLE OF CONTENTS

LIST OF TABLES

# CHAPTER 1 – Introduction

"A lot of people say I'm clever. I like it the way I am and I would like to stay clever. If you are clever it does not mean you are popular, because I am clever, but I am not popular."

–Nyla (Renold & Allen, 2006)

Question: "Have you heard the word "brainiac" used here?

Answer: "Yes. [When referring to students who take the Advanced Placement courses here.] That's a term for the smartest person in class. Brainiac--jerk--you know, those terms. If you're smart, you're a jerk, you're a brainiac."

Question: "So it's not a positive [term]?"

Answer: "No, it's a negative [term], as far as brilliant academic students are concerned."

–Sidney (Forham & Ogbu, 1986)

The two students who supplied these quotes could hardly be more different. At the time of her statement, Nyla was a 5$^{th}$ grade girl living and attending school in Southern Ireland in the early 2000's. She described herself as Pakistani, a practicing Muslim, and attended a school with students from a broad range of racial/ethnic and religious backgrounds. Nyla was at the top of her class and worked hard to achieve her academic success. However, Nyla also described being picked on and socially isolated. She was not shy about her academic skills and her confidence seemed to annoy her peers. Others described Nyla as "man-like," "a nightmare," "weird," and "mad." Nyla valued being "clever," but recognized that it took a significant social toll (Renold & Allen, 2006).

Sidney, on the other hand, was an African-American boy attending high school in an inner-city Washington, D.C. school, growing up nearly 20 years earlier in the 1980's. While Nyla worked hard gain academic recognition, Sidney appeared to actively suppress his scholastic achievement in order to avoid the social repercussions. Earlier in his school career, Sidney had achieved higher standardized achievement test scores than almost all of his peers and typically brought home good grades. In high school though, Sidney feared the social isolation that accompanied academic success and chose to underperform academically. As a result of his underperformance, Sidney hoped that he had never "given [his peers] a reason to" dub him a "brainiac" (Fordham & Ogbu, 1986, p. 188).

Despite the differences between these students' era, nationality, racial, religious, gender, and academic success differences, both seemed to have intimate knowledge of the struggles and social dangers associated with being academically motivated and successful in societies where female and minority students are expected to be neither. The social abuse that these students reported – social isolation by Nyla and verbal harassment by Sidney – both paint an all too familiar picture of peer bullying victimization that occurs throughout the education system.

## Rates of Bullying

As the stories of these two students illustrate, school bullying is a pervasive problem that has been shown to affect students of all ages, cultures, and academic ability levels (Smith & Brain, 2000). Despite the growing interest and intervention in both the community and in school settings, students are still being negatively targeted by their peers and exposed to bullying victimization at an alarming rate (Cook, Williams, Guerra, Kim, & Sadek, 2010; Field, Kolbert, Crothers, & Hughes, 2009; Fonagy, Twemlow, Vernberg, Sacco, & Little, 2005).

While the rates of bullying victimization can differ based on the population studied and methods used, nationally representative studies completed in the last 15 years indicate that between 10% and 50% of school-age children have experienced some level of bullying victimization (Bradshaw, Sawyer, & O'Brennan, 2007; Nansel et al., 2001; Robers, Zhang, Truman, & Snyder, 2012). The rates of bullying victimization, which have either remained constant or have increased over the 15 years of study, show that our nation's students are at a high risk of experiencing these aggressive interactions with their peers that are associated with many negative outcomes.

**Definition of Bullying**

In order to better understand bullying and to keep conceptualizations of the behaviors that qualify as bullying consistent, a three-part definition of bullying has been developed and is widely accepted in the literature (Gottheil & Dubow, 2001). First, bullying is considered to be a form of proactive or instrumental aggression (Brown & Parsons, 1998; Espelage & Swearer, 2003; Roland & Idsøe, 2001), which means that bullying behaviors are those that accomplish some goal for the aggressive actor and are not, typically, an immediate reaction to some provocation. The goals demonstrated by a perpetrator of bullying often appear to be related to his or her standing in his or her peer group and bullying behaviors are used to display dominance and power within that peer group (Salmivalli, 2010).

Second, for an interaction to be considered bullying, the behaviors need to be repeated over time (Monks & Smith, 2006; Nansel et al., 2001; Smith, Cowie, Olafsson,

& Liefooghe, 2002). This repetitious aspect of bullying creates a situation whereby the victim of bullying experiences prolonged feelings of fear, which can lead to one or more negative outcomes. Third, a power differential must exist between the bully and the victim (Espelage & Swearer, 2003). A power differential can take many different forms including differences in physical stature, differences in social power, or the presence of special knowledge that the bully can use to harm his or her victim (Sutton, Smith, & Swettenham, 1999). This three-part definition of bullying illustrates the specific types of interactions between two parties that should be considered instances of bullying, and paints a troubling picture of a potentially helpless victim who is repeatedly attacked by a more powerful bully who victimizes the target individual until the bully's goals are accomplished.

## Types of Bullying Behaviors

While the bullying definition described above illustrates what constitutes bullying behaviors in a general sense, this definition does not clearly operationalize the specific behaviors or tactics that an aggressive individual may use to bully others. Indeed, there are many different patterns of behavior that a bully could choose to use to harm his or her desired victim in the attempt to gain social dominance. These possible behaviors are most broadly divided into two groups – direct and indirect aggression.

As indicated by the terms used, direct aggression occurs when a bully confronts his or her victim face-to-face and commits the aggressive act (Richardson & Green, 2006). Direct aggression can be further delineated into two distinct categories – physical aggression and verbal aggression. Physical aggression includes such actions as hitting, pushing, slapping, biting, and kicking, which can be used to physically harm or intimidate the victim (Card, Stucky, Sawalani, & Little, 2008; Griffin & Gross, 2004). Despite the visibility of physical bullying, national studies show that between 9% and 13% of the student population is harassed in this fashion (Robers et al., 2012; Wang, Iannotti, & Nansel, 2009).

Verbal aggression includes any type of aggressive or derogatory communication made directly to the victim, including statements that mock the victim's appearance, intelligence, physical abilities, or any other topic that calls attention to a power differential between the bully and the victim (Camodeca & Goosens, 2005; Griffin & Gross, 2004). Research suggests that verbal aggression is the type of aggression that is most commonly reported by victims of bullying, and between 19% and 31.5% of students nationwide have been bullied via verbal aggression (Robers et al., 2012; Wang et al., 2009).

Unlike direct aggression where the bully and victim are both present in a single space and time, and, therefore, the victim is immediately aware of the aggressive act being directed towards him or her, indirect aggression is generally more subtle and involves attempts aimed at harming the victim's social status, relationships, or property (Richardson & Green, 2006). While there appears to be general consensus regarding the presence of indirect bullying behaviors, there is some debate over the specific terms and definitions that should be used when describing these aggressive behaviors. Researchers have used different labels, including indirect aggression, social aggression, and relational aggression to describe the different behaviors and aggressor motivations that could be applied during bullying relationships (Crothers, Schreiber, Field, & Kolbert, 2009; Spears, Slee, Owens, & Johnson, 2009). While these differing conceptualizations of indirect aggression make the determination of the rates of the behaviors difficult, studies examining rates of aggressive behaviors suggest that roughly 25% of bullying victims experience indirectly aggressive behaviors (Wang et al., 2009).

## Negative Outcomes of Bullying

Not surprisingly, students who experience victimization from any of these types of aggressive behaviors are at an increased risk of experiencing negative outcomes in their everyday lives. Previous research has demonstrated that bullying victimization exists in groups of Kindergarten students, and persists throughout students' school experience and on into adulthood (Kochenderfer & Ladd, 1996; Lipinski & Crothers, 2014). Students who experience bullying victimization have been shown to experience increased rates of mental health problems, including depression, anxiety, suicidal ideation, suicide attempts, loneliness, and lower levels of self-esteem as compared to their non-bullied peers (Craig, 1998; Fekkes, Pijpers, Fredriks, Vogels, & Verloove-Vanhorick, 2006; Klomek, Marrocco, Kleinman, Schonfeld, & Gould, 2008; Kochenderfer & Ladd, 1996; Seals & Young, 2003). Bullied students have also been more likely to demonstrate increased somatic complaints, which may include feeling tense, poor appetite, bedwetting, abdominal pain, sleeping problems, feeling tired, vision problems, dizziness, digestive problems, difficulty breathing, somatic disorders, and skin conditions than their non-bullied peers (Fekkes et al., 2006; Houbre, Tarquinio, Thuillier, & Hergott, 2006). Perhaps most startling, a study has shown that almost 66% of students who have been bullied reported an elevated rate of negative symptomology that is consistent with a Post-Traumatic Stress Disorder (PTSD) diagnosis (Houbre et al., 2006).

Given all of these negative mental and physical health outcomes experienced by students who are bullied, it is no surprise that these students would likely struggle to succeed academically.  Research has shown that victims of bullying are more likely to be absent from school, view their school as a dangerous, unsupportive place, and experience lower levels of academic achievement compared to their non-bullied peers (Glew, Fan, Katon, & Rivara, 2008; Juvonen, Nishina, & Graham, 2000; Kochenderfer & Ladd, 1996; Wienke Totura, Green, Karver, & Gesten, 2009).  While most studies generally examine the level of functioning shown by students after they experience bullying victimization, one investigation found that past experiences of bullying predicted future academic success in victimized students; suggesting that students who experience bullying victimization show lasting negative academic achievement-related outcomes (Schwartz, Gorman, Nakamoto, & Toblin, 2005).

## Factors Associated with Bullying

Research has consistently shown the negative effects of bullying victimization. Due to these negative outcomes, many researchers and school personnel have attempted to uncover which students are at the greatest risk for becoming victims of bullying.  As a result of these investigations, several student characteristics have been found to be associated with student victimization.

One of the most consistent predictors of bullying victimization is the student's age.  As noted, bullying has been shown to exist across all levels of development, from preschool and kindergarten (Crick, Casas, & Ku, 1999; Kochenderfer & Ladd, 1996) through high school (Li, 2010) and even on into adulthood (Lipinski & Crothers, 2014). While bullying does appear to take place across all age ranges, representative, nationwide surveys of bullying behavior have shown that in school-age children, bullying victimization rates appear to increase as children age, peaking in the late elementary to early middle school years. Bullying behaviors are then reported to dip slightly as students move through high school (Nansel et al., 2001; Robers et al., 2012).

Students' race is also frequently noted as a factor associated with a variety of differences in students' social experiences, including rates of bullying victimization. Studies have consistently shown that White students report bullying victimization most frequently when compared to racial/ethnic minority peers, including Black, Hispanic, and Asian students (Hanish & Guerra, 2000; Peguero, Popp, & Koo, 2011; Williams & Peguero, 2013). There is some debate, however, regarding the accuracy of these reports, as minority students' conceptualization of bullying in association with established cultural norms may lead minority students to underreport their bullying experiences (Sawyer, Bradshaw, & O'Brennan, 2008).

Students' gender has typically been described as another predictor of bullying victimization, with males showing higher rates of aggressive behavior than females (Björkqvist, Lagerspetz, & Kaukiainen, 1992; Crick, Bigbee, & Howes, 1996). But contrary to this popular belief, recent research has begun to suggest that the link between a student's gender and completing or experiencing aggressive acts is less certain (Archer & Coyne, 2005). Studies using large student samples have shown that girls may actually experience and perpetrate the same, or even more, aggressive behaviors than boys (Robers et al., 2012; Tulloch, 1995). One gender-related characteristic, however, still appears to hold: boys complete and experience more direct aggression while girls are more likely to participate in indirect aggression (Card et al., 2008; Lagerspetz, Björkqvist, & Peltonen, 1988; Wang et al., 2009).

Similarly, popular beliefs about bullying victimization suggest that academically successful students are more likely to be bullied. While there is research to support this sentiment (e.g., Bishop et al., 2004; Brown & Steinberg, 1990), some studies have shown that academically gifted students may actually be less likely to be victimized (Estell et al., 2009; Peguero et al., 2011). Some of this research indicates that academically successful students are frequently rated as possessing the pro-social and problem solving skills necessary to avoid or work through social situations where bullying may occur (Janke & Lee, 1991; Preuss & Dubow, 2004). Interestingly, however, while academic success may not be associated with increased rates of bullying victimization, academic effort and interest may be related to an increased risk. Some studies have shown that students who participate in school-related and academics-related activities such as honor societies, band, student government, or extracurricular clubs such as academic or service groups report elevated rates of victimization (Peguero, 2009; Popp & Peguero, 2011). These findings may suggest that academic interest or effort is a greater risk factor than academic success in general.

Perhaps the most consistent risk factor for bullying victimization is being different in some way from the student's peer group. Students who are less physically attractive, overweight, display a physical or performance disability, or display non-stereotypic gender behaviors are all more likely to experience increased rates of bullying victimization (Sweeting & West, 2001; Young & Sweeting, 2004). Furthermore, these differences are not independent of one another and students who display two or more of these differences show even higher rates of victimization (Sweeting & West, 2001).

## Expectations of Academic Achievement

One way that students can behave differently than the majority of their peer group is to demonstrate levels of academic interest and success that is typically uncommon from their social group. There is a long, documented history of racial/ethnic minority groups, as well as White female students, demonstrating lower levels of academic achievement than their White and male peers (Coleman et al., 1966; Ladson-Billings, 2006). This difference in academic achievement has been popularly titled “the achievement gap” (Ladson-Billings, 2006).

Despite increased awareness and attempted interventions to help reduce this gap, the differences in overall scholastic achievement between White students and minority students have largely remained (Barone, 2011; Good, Aronson, & Harder, 2008; Lee, 2002). Research still shows that on average, Black and Hispanic students lag behind White and Asian students in every academic category, while girls perform noticeably worse in the areas of math and the physical sciences when compared to boys (Aronson, Quinn, & Spencer, 1998).

Perhaps even more troubling than the achievement gap is what Plucker, Burroughs, and Song (2010) refer to as the “excellence gap.” These researchers found that governmental initiatives like No Child Left Behind (NCLB) were at least somewhat effective at reducing the achievement gap by increasing the number of minority students who were competent in reading and math. But as the overall achievement gap began to shrink, the gap in the level of academic achievement between the highest achieving

White students and the highest achieving minority students, and between the highest achieving male and highest achieving female students, has widened (Plucker et al., 2010). This widening of the excellence gap is likely to ensure that minority and female students will still face the same glass ceilings of educational and career achievement that have plagued these groups for years.

## Math and Science Achievement

While the achievement gap found broadly between different groups' academic functioning has garnered a large amount of interest, research, and intervention over the past 50 years, a more specific area of academics has begun to receive more attention. Government and private agencies have begun tracking and advocating for increased educational standards and opportunities in the areas of math and the physical sciences. Often, these initiatives that promote math and science education are broadened to include technology and engineering coursework, and this set of courses and educational topics are frequently considered together and simply referred to as the Science, Technology, Engineering, and Math, or STEM, fields. Broadly, STEM fields are defined as those areas of education pertaining to mathematics, the natural sciences, physical sciences, biological/agricultural sciences, engineering/engineering technologies, computer/information sciences, and the social sciences (Chen & Weko, 2009; United States Government Accountability Office, 2012).

The increased interest in these fields is largely due to the realization that the U.S. education system has slipped in its ability to produce competent and capable students in math and science areas. When compared to 34 other industrialized nations on a standardized mathematics and science assessment, the United States was found to rank

25th in students' mathematics achievement and 17th in science achievement (United States Department of Education, 2013). Furthermore, of the high school seniors who are interested in entering a STEM-related field, only 16% have been found to be proficient in math (United States Department of Education, 2013).

As a result, the U.S. government invested $3.1 billion to fund 209 different programs that directly supported STEM education (United States Government Accountability Office, 2012). One of the most commonly referenced goals of this investment of public resources is to reduce the achievement, participation, and interest gaps in STEM fields between White male students and racial/ethnic minority as well as White females students (United States Department of Education, 2013). White males have dominated STEM-related training programs and careers for many years (Crowley, 1977; Sells, 1976) and academic indicators such as student course selection and participation in Advanced Placement (AP) exams all show that White males are still attempting and achieving at rates much higher than their minority and female peers (The College Board, 2013; Dalton, Ingels, Downing, & Bozick, 2007).

**Why the Gap in Math and Science Achievement?**

Like the achievement gap at large, the math and science achievement and participation gaps between minority students and female students has been a topic for debate for some time and several different theories have been discussed. Some theories posit that the gap is a reflection of some cognitive deficit inherent to minority and female students (e.g., Deary, Penke, & Johnson, 2010; Roth, Bevier, Bobko, Switzer, & Tyler, 2001; Rushton & Jensen, 2005; Spelke, 2005; Summers, 2005). However, recent studies that have controlled for important social status, cultural value, or personal expectation variables have been able to equalize academic performance between White males and their minority and female peers (Cohen, Garcia, Apfel, & Master, 2006; Hyde, Lindberg, Linn, Ellis, & Williams, 2008; Steele & Aronson, 1995).

Rather than focusing on individual minority and/or female students' abilities or reactions to math and science tasks, another promising area of research has examined how a student's environment reacts toward a student who has demonstrated a level of academic achievement that is different from his or her social group's historical pattern of achievement. Qualitative researchers have found that when students behave differently than what is commonly expected in terms of academic effort and achievement, these students are often described as "acting White" by their peers (Fordham & Ogbu, 1986). By participating in what has traditionally been deemed "White" behaviors, such as "speaking standard English" and "getting good grades in school" (Fordham & Ogbu, 1986, p. 11), peers of minority students view the offending student as essentially rejecting their racial/ethnic heritage. Instead, the offending student is regarded as placing more importance in the traditional values of the dominant White culture. Those students who choose to act in these "White" ways consistently experience social reprimands that would certainly qualify as episodes of bullying victimization (Fordham & Ogbu, 1986). This "acting White" phenomenon has been shown to exist across time, geographic location, and socio-economic status (Fordham & Ogbu, 1986; Ogbu, 2003). Similar social reprimands have been shown to be levied on girls who violate traditional gender stereotypes of academic achievement and effort (Renold & Allen, 2006).

Other studies have been able to quantify this interaction between student minority or female status and academic success and its effects on student popularity. Fryer (2006) wanted to investigate what he termed as the "pressure to be average." In order to study this phenomenon, students were asked to supply a list of their closest friends, which, when aggregated amongst the whole sample, indicated students' level of popularity. Students were also asked to report their most recent academic grades, which were compared to their peer-described level of popularity.

Results showed that for White students, increased academic achievement was associated with greater popularity; the highest achieving students were generally rated as having the most friends. Results for minority students, however, differed noticeably. For Black students, increased academic achievement was slightly associated with increased popularity. Students achieving a grade point average (GPA) of 1.0 were rated as the least popular and ratings of popularity peaked for students who earned a GPA of 3.5. But when students achieved a GPA higher than 3.5, their reported level of popularity decreased as their GPA approached 4.0.

Findings for Hispanic students were even more startling. Again, increased GPA was associated with increased popularity to an extent, as students with GPAs of 1.0 were rated as being less popular than students with a GPA of approximately 2.5. Interestingly, however, the reported popularity of a student plummeted when he or she was found to achieve a GPA higher than 2.5. Popularity of these higher achieving Hispanic students fell to such a degree that students who achieve a 4.0 GPA were rated as having significantly fewer friends than even those students who earned a GPA of 1.0 (Fryer, 2006).

Whether described as "acting White" or the "pressure to be average," academic achievement appears to be related to negative social outcomes for those racial minority students attempting to excel. Therefore, minority students appear to have an incentive to not pursue academic achievement or success if they value their social relationships. This research suggests that students may choose to underperform in academic areas, including math and science, to maintain their social standing.

While none of the studies presented discussed math or science achievement specifically, these studies illustrate the problems experienced by students who attempt advanced math- and science-related classes. Math and science education comes loaded with expectations about who should, and should not, be able to achieve, and when students contradict those expectations and behave in an unexpected way, others often struggle to adjust to these violations. In response, those individuals appear likely to try to force the individual back to conformity (Prentice & Carranza, 2002).

### Bullying of Females and Minorities Who Are Academically Different

While the studies discussed above show peer responses that would qualify as bullying, the researchers completing the investigations did not examine bullying behaviors specifically. Other research, however, has examined the experience of bullying victimization for students who violate gender- or racially-based expectations. Two studies bypassed victims' reports of bullying and attempted to study potential aggressors' behaviors directly. These studies placed seemingly normal individuals in situations where they experienced another person either performing in a way that conformed to gender or racial expectations and then in situations where another person's behavior violated those expectations.

In both cases, individuals were significantly more likely to punish, sabotage, and undermine the future success of those people who violated gender and racial stereotypes than they were for those who conformed to stereotypes (Phelan & Rudman, 2010; Rudman & Fairchild, 2004). These findings show that those individuals who display skills and interest outside of those that are allowed by a society's traditionally held thoughts and beliefs are more likely to experience negative responses from those around them. Therefore, by extension, these results suggest that students who violate the math and science achievement expectations may experience active resistance by others in their attempts to succeed in math and science coursework.

Other investigations that have used nation-wide, representative data show that Black students who displayed academic success were significantly more likely to experience bullying victimization than White students who had the same academic success. The same trend was found with Hispanic students, although the likelihood of bullying did not reach significance (Peguero & Williams, 2011; Williams & Peguero, 2013). Similarly, Black, Hispanic, and Asian students who participated in academic extracurricular activities (e.g., academic honor society, service organization) were significantly more likely to experience violent victimization than White students who participated in the same types of activities (Peguero et al., 2011). The authors of these studies claim that these findings are further evidence that acting different from what is expected increases the risk for bullying victimization. Black students, and to some degree Hispanic students, are at a greater risk of bullying victimization when they display high academic achievement and higher rates of academic interest.

## Problem Statement and Research Questions

Even though this body of research does not mention math and science involvement specifically, it is clear that the characteristics of a student who in skilled and interested in math and science coursework are also associated with an increased risk of peer victimization for those students who violate the traditional beliefs which students should be participating in math and science education. Therefore, in order to help explain why the achievement gap in math and science education continues to persist despite increased governmental and societal interest and intervention, more information is needed to understand the social responses experienced by racial minority students and White female students who participate and succeed in math and science classwork during their early adolescent school years. In order to help gather more information regarding this topic, the following research questions and associated hypotheses were developed.

Research Question 1: Does a student's identification with a specific race impact his or her experience of peer bullying victimization after controlling for the student's ability, enjoyment, and confidence in mathematics?

> Hypothesis 1: After controlling for his or her mathematics ability, confidence, and enjoyment, a racial minority student will report higher frequencies of peer bullying victimization than a White student.

Research Question 2: Does a student's identification with a specific race impact his or her experience of peer bullying victimization after controlling for the student's ability, enjoyment, and confidence in science?

> Hypothesis 1: After controlling for his or her science ability, confidence, and enjoyment, a racial minority student will report higher frequencies of peer bullying victimization than a White student.

Research Question 3: Does a student's gender impact his or her experience of peer bullying victimization after controlling for the student's ability, enjoyment, and confidence in mathematics?

Hypothesis 1: After controlling for her mathematics ability, confidence, and enjoyment, a female student will report higher frequencies of peer bullying victimization than a male student.

Research Question 4: Does a student's gender impact his or her experience of peer bullying victimization after controlling for the student's ability, enjoyment, and confidence in science?

Hypothesis 1: After controlling for her science ability, confidence, and enjoyment, a female student will report higher frequencies of peer bullying victimization than a male student.

# CHAPTER 2 – Literature Review

## Bullying

School bullying is a pervasive problem throughout the educational system that has been shown to affect students of all ages, cultures, and ability levels (Smith & Brain, 2000). While the topic of bullying has received a growing amount of attention in the last decade, including investigations in the research literature (Cook et al., 2010), development of bullying intervention programs (Field et al., 2009; Fonagy et al., 2005), and government-led initiatives (Stuart-Cassel, Bell, & Springer, 2011), the problem persists.

While the frequency of bullying in schools has varied according to the samples studied, research that has investigated large and diverse samples of students suggests that bullying rates have remained stagnant or have even increased over the last 10 to 15 years. A nationally representative study completed in 1998 related to the episodes of bullying victimization experienced by $6^{th}$ through $10^{th}$ grade students found that 10.6% of students were at least "sometimes" victims of bullying by their peers, while a total of 29.9% of students were in some way involved in the bullying cycle at school (Nansel et al., 2001).

Eight years later, an investigation involving over 15,000 students gathered from a diverse group of geographically located, racial populated, and economically privileged schools found that nearly half (49%) of students in grades 4 through 12 reported being bullied at least once within the last month (Bradshaw et al., 2007). Finally, one of the most recent nationally representative studies commissioned by the U.S. Department of Education found that 28% of students between the ages of 12 and 18 reported being bullied in the last year (Robers et al., 2012). These findings suggest that nearly a third of the student population is experiencing bullying victimization at the hands of their peers and, therefore, are at an increased risk for psychological-, social adjustment-, and academic achievement-related problems.

**Definition of Bullying**

The first step in understanding the concept of bullying includes establishing behaviors and interactions that constitute an experience of bullying. In the broadest sense, bullying is considered a form of aggressive behavior. Aggressive behavior is typically defined as hurtful acts that are purposefully completed toward another individual in the hopes of harming that individual (Smith et al., 2002). This type of behavior would be in contrast to accidental behavior, in which the harm that one causes to another is not purposeful.

Aggression has been further divided into two broad categories: proactive and reactive aggression. Proactive aggression occurs when the aggressive individual takes some step to plan his or her behavior and concludes that acting in an aggressive manner would bring about some reward. Reactive aggression, conversely, occurs when the aggressive individual responds to an adverse stimulus in the environment and uses aggression as an immediate response to this stimulus (Brown & Parsons, 1998).

This qualification, however, conceptualizes aggressive actions from the aggressor's point of view and not from the victim's. Geurin and Hennessy (2002) argue that an act may be considered aggressive if the victim perceives the act as harmful even when the aggressor does not intend to do harm. This would broaden the number of interactions that could be considered aggression and, therefore, bullying by extension. While this is an important consideration, the majority of the research regarding aggression and bullying considers the bully's point of view and how his or her perceptions of his or her behavior qualify his or her actions as aggressive or not.

Since not all aggressive acts would be considered bullying, more specific distinctions are needed to determine when simple aggression becomes an episode of bullying. After years of research and competing theories, present day researchers have agreed on three general factors that are necessary to consider an aggressive act as bullying. First, bullying is considered a form of proactive or instrumental aggression that does harm to the victim (Brown & Parsons, 1998; Espelage & Swearer, 2003; Roland & Idsøe, 2001). This suggests that bullying behaviors are not typically aggressive acts in which individuals are responding to a provocation, but rather behaviors initiated by the bully that serve a purpose or help the bully to accomplish some end. A bully's goals typically appear related to his or her standing in his or her peer group and bullying behaviors are used to display dominance and power within that peer group (Salmivalli, 2010).

Second, bullying behaviors are repeated over time (Monks & Smith, 2006; Nansel et al., 2001; Smith et al., 2002). Some claim that a single instance of purposeful aggression can still be considered bullying if that instance creates a prolonged feeling of fear and worry in the victim (Geurin & Hennessey, 2002), but most researchers indicate that true experiences of bullying occur when aggressive behaviors continue repeatedly. This repetitious aspect of bullying creates a situation whereby the victim of the bullying experiences prolonged feelings of fear, which lead to one or more negative outcomes.

The third component of the bullying definition indicates that, in the bullying relationship, a power differential must exist between the bully and the victim (Espelage &

Swearer, 2003). Power differentials may be related to differences in physical stature between the bully and the victim. A bully may also wield more social power and already have a higher social standing than his or her victim. Or, a bully may hold special knowledge that the he or she can use to harm his or her victim (Sutton et al., 1999). This three-part, or tripartite definition of bullying (Gottheil & Dubow, 2001), illustrates what specific types of interactions between two parties should be considered instances of bullying. The definition also paints a troubling picture of a potentially helpless victim who is repeatedly attacked by a more powerful bully who victimizes the target individual until the bully's goals are accomplished.

**Types of Bullying Behaviors**

While defining what types of interactions qualify as bullying is an important part of confronting the problem, the established definition does not clearly operationalize what a specific episode of bullying may look like. Indeed, there are many different patterns of behavior that a bully could use to harm his or her desired victim in the attempt to gain social dominance. Such behaviors are most broadly divided into two groups – direct and indirect aggression. As indicated by the terms used, direct aggression occurs when a bully confronts his or her victim face-to-face and commits the aggressive act (Richardson & Green, 2006). In this case, the victim knows immediately that he or she is being bullied and also know exactly who is responsible for causing the distress.

The concept of direct aggression has been further delineated into two specific types of behaviors that may be exhibited. One specific domain of behaviors is known as physical aggression. Physical aggression includes such actions as hitting, pushing, slapping, biting, and kicking which can be used to physically harm or intimidate the victim (Card et al., 2008; Griffin & Gross, 2004). These types of behaviors are often difficult to conceal, and both peers and adults in close proximity to the victim often witness these aggressive acts. The visibility of these physically aggressive behaviors likely contributes to the relatively low percentage of students, between 9% and 13% of the population, who report being harassed in this fashion (Robers et al., 2012; Wang et al., 2009).

Another form of direct aggression is known as verbal aggression. This type of aggression includes any type of aggressive or derogatory communication made directly to the victim, including statements that mock the victim's appearance, intelligence, abilities, or any other topic that calls attention to a power differential between the bully and the victim (Camodeca & Goosens, 2005; Griffin & Gross, 2004). Verbal aggression has been found to be the most frequently utilized type of aggression employed by aggressive students. Results of recent nationwide investigations indicate that between 19% and 31.5% of students have reported being verbally harassed (Robers et al., 2012; Wang et al., 2009).

While direct bullying typically occurs in the presence of the victim, indirect aggression is generally more subtle. Those bullies who choose to harm their victims in an indirect manner do so without the victim's immediate knowledge, often leaving the victim unsure, or at least lacking proof, of which person is the cause of his or her distress (Björkqvist et al., 1992). This method of bullying occurs when the bully attempts to cause harm by damaging his or her victim's social status, relationships, or property (Richardson & Green, 2006).

While some researchers consider the term, indirect aggression, to be a sufficient label in describing these bullying behaviors that are completed without the victim's full knowledge and understanding (Lagerspetz et al., 1988; Richardson & Green, 2006), others feel as though indirect aggression encompasses several different constructs, namely, constructs known as relational and social aggression (Crick & Grotpeter, 1995; Galen & Underwood, 1997). Recent research has provided evidence that these separate constructs do explain important behavioral and motivational differences in bullies' actions and, therefore, deserves independent consideration (Crothers et al., 2009; Spears et al., 2009).

The primary distinction between these two constructs is the means by which the bully manipulates the victim's peer relations in order to cause harm or force compliance. In relational aggression or bullying, the bully uses his or her relationship with the victim to cause harm. This generally involves the bully threatening to affect the victim's individual relationships either with the bully or with other peers in order to force the victim to comply with his or her desires (Archer & Coyne, 2005; Crick & Grotpeter, 1995). Relational bullying can occur both directly (e.g., threatening to not be the victim's friend, physically avoiding the victim) or indirectly (e.g., gossip, cruel rumors; Cullerton-Sen & Crick, 2005). Relational bullying appears to occur at a similar rate as verbal bullying, as more than 25% of students reported being involved in relational bullying during the last year (Wang et al., 2009).

In social aggression or bullying, however, the bully does not affect individual relationships but instead causes harm by affecting the victim's standing in the peer group at large. A socially aggressive bully uses his or her popularity with peers to convince others to avoid the victim or spreads rumors that will affect the victim's social standing, thereby using the whole peer group to do harm to the victim (Archer & Coyne, 2005; Richardson & Green, 2006). This style of bullying almost always occurs indirectly and takes the victim of the aggression by surprise. Due to the indirect manner of the bullying experience, the victim is likely to feel especially demoralized and unable to fight back against the bully as he or she has lost his or her social support and may not even know which individual is responsible for instigating the aggression against him or her (Xie, Swift, Cairns, & Cairns, 2002). Due to the indirect nature of social aggression, gathering specific rates of the behavior is difficult. But in a study examining the specific behaviors students experienced during episodes of bullying, indirect behaviors consistent with social aggression (e.g., gossiping, making fun of others behind their back, getting others to not like someone) occurred with the second highest frequency, just slightly behind the frequency of verbal aggression (Coyne, Archer, & Eslea, 2006).

Recent research has also begun to explore the phenomenon of cyberbullying (Li, 2006; Spears et al., 2009). As research into the topic grows, different theories have arisen about how to classify cyberbullying. Beran and Li (2007) conceptualize cyberbullying as a new method of perpetrating indirect aggression against a targeted victim. These researchers compare cyberbullying to the definition of indirect aggression posited by Björkqvist et al. (1992) that discusses aggression done without the victim's immediate knowledge while targeting the victim's peers and social standing to cause the victim harm. According to Beran and Li (2007), cyberbullying matches this definition, as the internet often provides the bully with anonymity while still providing the access to both the victim and the peer group necessary to demonstrate the power differential between the bully and victim.

Spears and her colleagues (2009), however, wrote that cyberbullying deserves its own separate consideration and these researchers coined the term “covert bullying” to explain the cyberbullying phenomenon. This group claimed that the construct of covert bullying encompasses the ideas of indirect, social, and relational aggression as they felt that cyberbullying provides bullies with access to all the behaviors and motivations described in the previously discussed terms. Additionally, covert bullying includes the extreme anonymity that cyberbullying allows as students hide behind screen names, access to others’ information and accounts, and the ability to secretly spread personal information with little or no repercussions (Spears et al., 2009).

The rates at which cyberbullying occurs can also be difficult to determine. Li (2010) asked students about their responses to cyberbullying and found that 42.5% of 7$^{th}$ through 12$^{th}$ grade students would “Do nothing” about the victimization they were experiencing, and only 11.7% of students would tell an adult. This suggests that students who experience cyberbullying may not be accurately reporting their experience, which impacts researchers’ ability to accurately study the cyberbullying phenomenon. This also impacts the ability to create effective interventions that target cyberbullying aggressors and victims.

Not surprisingly, studies that have attempted to determine how frequently cyberbullying occurs have yielded varying rates. Robers et al. (2012) conducted a survey of 12- to 18-year-old United States students in which they found that 6% of the student population reported experiencing cyberbullying. Wang et al. (2009) found slightly higher rates in their analysis of a nationwide data set, as they concluded that at least 8.1% of students experienced some form of cyberbullying. However, studies that have used samples from a specific group or school district have found significantly higher rates of cyberbullying. In her study of 264 Canadian 7th through 9th grade students, Li (2006) found that roughly 25% of students reported experiencing cyberbullying. Similarly, Erdur-Baker's (2010) investigation reported that 22.5% of 276 Turkish 14- to 18-year-olds reported being victimized through cyberbullying. These rates of occurrence are similar to those rates of other, more frequently studied, forms of bullying and suggest that cyberbullying is a significant problem and growing concern for students and schools.

**Negative Effects of Bullying**

Students who are bullied during school have consistently shown a large number of negative effects related to this experience of victimization. These negative effects can start as early as kindergarten. Kochenderfer and Ladd (1996) studied a group of 200 kindergarten students and found that those students who reported being bullied were significantly more likely to report feeling lonely and make attempts to avoid school. These researchers also concluded that, because students demonstrated a desire to avoid school only after the initiation of bullying victimization, there was evidence of a causal link between bullying victimization and school avoidance. Furthermore, these negative feelings toward school, a desire for social avoidance, loneliness, as well as feelings of depression may arise only months after a student begins to be chronically victimized and can continue for many years, even if the bullying victimization stops (Juvonen et al., 2000; Kochenderfer & Ladd, 1996). Besides depression and loneliness, victims of peer bullying are also at risk for such mental health problems as anxiety, serious suicide ideation, suicide attempts, and lower levels of self-esteem as compared to their non-bullied peers (Craig, 1998; Fekkes et al., 2006; Klomek et al., 2008; Seals & Young, 2003).

In addition to increased mental health risks, victims of bullying may also suffer from significant somatic complaints. Fekkes and colleagues (2006) asked 1,118 Norwegian 9- to 11-year-old students about their status as a victim of bullying and their health complaints at the beginning and end of a school year. Results showed that students who had begun experiencing bullying victimization during that school year were also significantly more likely to develop somatic complaints by the end of the year. These newly bullied students complained of such somatic symptoms as feeling tense, poor appetite, bedwetting, abdominal pain, problems sleeping, and feeling tired (Fekkes et al., 2006).

Houbre and colleagues (2006) found similar results in a smaller French sample and concluded that victims of bullying experienced higher rates of vision problems, dizziness, digestive problems, difficulty breathing, somatic disorders, and skin conditions. This research team also studied if these students' experiences of bullying led to symptoms consistent with Post-Traumatic Stress Disorder (PTSD) as measured by the Impact Event Scale (Horowitz, 1979). They found that, of the students who reported being bullied one or more times in the last year, 65.8% evidenced a high post-traumatic stress level, while only 8.2% reported a low level (Houbre et al., 2006). This finding suggests that students who are bullied are at risk of developing serious mental and physical health problems.

With these significant mental and physical health issues, students who are bullied often struggle to succeed at school. Research has shown that bullying victimization is associated with victims' reports of lowered liking of school and lowered academic success, while reporting higher rates of school absenteeism (Juvonen et al., 2000; Kochenderfer & Ladd, 1996; Wienke Totura et al., 2009). Experiencing bullying may also lead to a significant increase of students feeling unsafe at school and perceiving that they do not fit in or belong at the school that they attend (Glew et al., 2008).

Also, students who experience bullying at school are less likely to be able to focus on their academic work as their anxiety regarding future victimization distracts them from learning (Card & Hodges, 2008). Schwartz, Gorman, Nakamoto, and Toblin (2005) found that in a sample of 3rd and 4th grade students, past bullying victimization predicted future academic success. The researchers of this study concluded that students' preoccupation with their experiences of victimization distracted them from achieving appropriately at school.

Overall, victims of bullying appear to be a group of students who are at risk for developing significant mental and physical health problems. These students tend to be more fearful and avoidant of school, and even when they do attend school, these students demonstrate a lower level of academic performance than they may have achieved had they not experienced periods of bullying victimization.

**Student Characteristics and the Relationship with Bullying Experience**

**Age.** Bullying has been studied in populations ranging from students in preschool and kindergarten (Crick et al., 1999; Kochenderfer & Ladd, 1996) through high school (Li, 2010) and even on into adulthood (Lipinski & Crothers, 2013). While bullying does appear to occur across all ages, the rate at which bullying victimization occurs at each different point in development is debated in the literature.

Based on their meta-analysis of 20 years of bullying research that examined the behavior of students ranging in age from 3 to 18, Cook and colleagues (2010) found that the age of the student was not related to any significant change in the rate of bullying victimization. However, other studies that gathered specific rates of bullying from large samples have demonstrated that the rate at which students report experiencing bullying does vary with age. The report by Nansel et al. (2001) indicated that students report the highest frequency of bullying (e.g., "Sometimes" or "Weekly") in the 6th grade, as 24.2% of students report being victimized. In this study, the reported frequency of bullying decreased consistently through students' 10th grade year, when 9.4% of students reported experiencing bullying victimization. Robers et al. (2012) reported a very similar trend, although with higher overall rates. This study found that 39.4% of 6th graders and 22.2% of 12th graders reported bullying victimization.

One theory that attempts to explain this reduction in bullying rates posits that the actual rate of bullying does not change with age; rather, what changes is students' understanding and conceptualization of what should be considered bullying behavior. In order to better understand students' conceptualizations of the types of behaviors that constitute bullying, two studies administered drawings and short vignettes that illustrated different social interactions to sets of elementary-aged and middle school-aged students. The researchers then asked students if they considered what they had just read to be an episode of bullying (Monks & Smith, 2006; Smith et al., 2002). Across demographic classifications, 6- to 8-year-old students consistently classified all aggressive behavior as bullying; including behaviors that would not meet the definition of bullying. These students' classification of bullying was also very simplistic and did not consider there to be a difference between direct or indirect types of aggression. This pattern of classification led to over-reporting of bullying behavior.

Fourteen-year-old adolescents, however, were able to make more accurate and nuanced classifications of behaviors that would qualify as bullying and were also more likely to differentiate between different types of behaviors that bullies may use. These older students were able to respond in a fashion much more consistent with both the tripartite bullying definition and with adult classifications of behaviors. As a result, while the adolescents and elementary-aged students were exposed to the same pictures and vignettes, adolescents classified fewer of the interactions as bullying. Smith et al. (2002) and Monks and Smith (2006) theorized that it is this increased accuracy in identifying which interactions should, or should not, be considered bullying that leads to the apparent reduction of bullying behaviors as students age, not actual reductions in bullying behaviors.

These theories suggest that while bullying rates appear to fall as children age, these rates may not reflect actual behaviors, but rather indicate the growing cognitive sophistication and understanding in older students. The cause of the reduction in reports of bullying is due to children having a better understanding of bullying behavior, gaining the ability to differentiate types of behaviors, and being able to better understand the motivations of the bully. Therefore, older students' reports are likely a more accurate picture of the rate of bullying that is occurring while younger students may be over-reporting bullying behaviors.

**Gender.** Traditionally, males have been considered to be more aggressive than females and, therefore, more likely to bully and be bullied than their female peers. In some ways, research has supported this notion. Boys have been found to display higher rates of direct aggression than girls, especially during early to middle childhood (Björkqvist et al., 1992; Crick et al., 1996). However, most research paints a complex relationship between gender and the likelihood of being bullied and to bully others (Archer & Coyne, 2005).

Overall rates of victimization by gender have also been shown to vary across studies. Nansel et al. (2001) reported that more boys (20.7%) than girls (13.7%) reported at least "Sometimes" being bullied at school. However, a more recent study commissioned by the U.S. Department of Education found that girls (29.5%) were more likely to experience bullying victimization than boys (26.6%; Robers et al., 2012). While Nansel and colleagues' report showed rather noticeable differences with males reporting higher rates of victimization, Robers and colleagues' investigation cites victimization rates between male and female students back to 2005, and each year females were shown to be more frequently victimized than their male peers.

Tulloch (1995) asked 883 $8^{th}$ grade students about their bullying experiences. This study found that, while a significantly higher proportion of boys than girls reported that they had perpetrated bullying against one or more of their peers, rates of reporting bullying victimization were not significantly different between the genders. Furthermore, male victims of bullying reported that other male students were most likely to act as the perpetrator of the bullying behaviors, while female victims report being bullied by both male and female bullies equally.

The study went on to delineate the gender differences in the types of bullying behaviors exhibited, and found that the manner in which females were most likely to be bullied varied by the gender of the aggressive student. Both male and female victims report being hit, pushed, picked on, or teased more often by male bullies. However, female victims were significantly more likely to be ignored or excluded, spread rumors about, or threatened by other females than males. These findings support the idea that, while males are more directly aggressive, females tend to use more indirect aggression, especially when targeting members of their same gender.

Studies of the relationship between gender and bullying have repeatedly reported mixed findings. In one meta-analytic investigation of 148 different studies, results showed that boys were significantly more likely to commit direct aggression than girls, while girls were shown to exhibit more indirect aggression than boys at a statistically significant rate. Authors, however, state that while the gender difference in indirect aggression was significant, the difference was trivial in magnitude (Card et al., 2008).

Similarly, in a study of 5$^{th}$ grade students, Lagerspetz et al. (1988) found that, while boys were more aggressive overall, girls were reported to use significantly more indirectly aggressive behaviors than boys. Wang et al. (2009) also found that noticeably more boys (9.9%) than girls (5.5%) experienced physical bullying victimization while more girls (27.6%) than boys (20.7%) experienced more relational bullying victimization. Overall, while these studies have yielded some conflicting results, research seems to suggest that there are gender differences in both the rates of bullying and victimization, as well as the types of bullying behaviors used.

**Race.** Several studies have examined how race and ethnicity are related to bullying. Many of these studies have found that differences do exist in the rates that bullying occurs for racial minority students. Perhaps surprisingly, these studies suggest that racial minority students frequently report lower rates of victimization than their White peers. Hanish and Guerra (2000) found in a large scale study that Latino 1st through 6th grade students were less likely to be bullied by their peers than were Black and White students. Peguero and colleagues (2011) supported this finding that Latino students were less likely to be bullied and added that Asian students were also significantly less likely to be bullied than their Black and White peers. Finally, Williams and Peguero (2013) reported that Latino (37%), Asian (35%), and Black (37%) students all experienced significantly lower rates of peer victimization than were reported by White (43%) students. However, some researchers claim that ethnic minority youth, especially Black youth, underreport their experiences of being bullied (Sawyer et al., 2008).

Based on the power differential requirement in the bullying definition, it may seem intuitively appropriate to assume that bullying is primarily completed by a racial or ethnic majority student, and, therefore, theoretically more socially powerful, against a minority student. However, this use of race as an inherent power differential may not be accurate. In fact, the most significant negative outcomes appear to be experienced by those victims who are targeted by their same-group peers. When two students from the same racial group engage in a period of bullying, the victim during that interaction has been shown to experience increased amounts of loneliness and social anxiety (Bellmore, Witkow, Graham, & Juvonen, 2004; Graham, Bellmore, Nishina, & Juvonen, 2009).

**Academic interest and success.** Bullying of academically interested and successful students is frequently depicted in American pop culture. The terms of "nerd" and "geek" are often used in ways that would qualify as verbal aggression and suggest that academic success makes an individual a target for bullying behavior. Some research does support this belief. A study by Bishop et al. (2004) found that honors students and students who took accelerated courses in middle school experienced higher rates of peer harassment than did students who did not apply themselves as consistently to their academic work. These researchers go on to mention that when a student's grade point average (GPA) becomes significantly different (either higher or lower) than the school average, that student is placed at greater risk for peer victimization. Qualitative patterns discussed in the research seem to suggest that academic effort, and not necessarily academic success or GPA, may be the most significant predictor of social rejection and peer harassment (Bishop et al., 2004). Brown and Steinberg (1990) go on to suggest that students actively underperform academically in order to avoid the social stigma that accompanies academic success. This research seems to suggest that those who succeed academically may be at a higher risk of being bullied by their peers.

Recent research, however, paints a slightly different picture. Academic success has been found to be an insulating factor for some students and may lead to a lower experience of school-based violence (Peguero et al., 2011). Other research has found that academically gifted students were rated as more socially desirable by their peers and experience lower rates of bullying victimization than other students in general education and students with mild disabilities (Estell et al., 2009). Academically gifted students are also more likely to use effective problem solving skills when dealing with social conflict

(Preuss & Dubow, 2004), and have been rated as having higher rates of positive social skills than their general education peers (Janke & Lee, 1991). These findings suggest that gifted students, as a group, are less likely to experience bullying victimization and should have the social skills and supports necessary to avoid these types of negative social interactions.

Still, Peterson and Ray (2006) found that 67% of academically gifted students reported being bullied at some point during their elementary and middle school (K-8) careers, a rate which is higher than some reports related to the experiences of the general student population described earlier (Bradshaw et al., 2007). But, while 67% of students report some instance of bullying, only 11% of gifted students, report being consistently bullied over time, which is typically an important component in the definition of true bullying victimization (Peterson & Ray, 2006).

A study by Popp and Peguero (2011) showed that male students who participate in extracurricular clubs, including clubs that are related to scholastic achievement, are more likely to experience peer victimization than male students who participate in athletic-related extracurricular activities. Participation in non-athletic extracurricular groups such as an academic-related honor society, band, or student government, as well as clubs such as academic or service groups have been shown to increase the likelihood of violent victimization for all students who participate (Peguero, 2009). These findings suggest that those students who are most likely to put in extra time and effort to learning-related activities are more likely to be bullied than those who do not participate in academic-related groups.

These reports indicate that, while the academically gifted students, as a whole, are more likely to be prepared and have the skills necessary to navigate the difficult social environment, they are still at risk for being bullied by their peers and may come to experience the negative impacts brought about by peer victimization.

**Victimization of those who are "different".** Sweeting and West (2001) discuss a large body of research that describes how students who are perceived as different from their peers are at greater risk for social difficulties and peer victimization. In their own study, these researchers concluded that those 11-year-old students who were rated as less physically attractive or more overweight, demonstrated a noticeable physical or performance disability, or performed poorly academically, were significantly more likely to experience peer victimization. Furthermore, these characteristics were found to be independent of each other and, therefore, students who display two or more of these characteristics were even more likely to experience negative outcomes (Sweeting & West, 2001).

A separate study examined whether or not adolescents conformed to traditional gender roles as they interacted with their peers. In this study, if the student failed to display behaviors that were consistent with traditional gender norms, the researchers sought to determine if that student reported an increased rate of peer victimization. In order to answer this question, researchers asked students to describe their typical behaviors in order to determine which males acted in a more feminine manner and which females engaged in more masculine behaviors. Results from over 2,000 Scottish adolescents found that more effeminate boys were bullied significantly more often than their male peers who displayed more gender consistent behaviors. Similarly, more masculine girls also reported higher rates of victimization; however, these rates did not reach statistical significance (Young & Sweeting, 2004).

Finally, Patterson and Bigler (2007) attempted to manipulate the phenomenon of having some students portray atypical group characteristics in order to gauge the social outcomes experienced by these atypical students. Researchers gained access to 97 five- to eleven-year-old students during a summer camp. During the camp, the group distributed shirts to each of the participating students; half received Kelly Green shirts and half received Royal Blue shirts. Classes were divided into rooms of 15 to 17 students by age and an even number of green and blue students were placed in each room. However, in each room, 2 to 3 students were given a shirt that was a lighter shade of each of the particular colors (light green and light blue) creating a minority of atypically dressed students in each classroom. Students were required to wear these shirts every day during the summer camp, which lasted for four weeks. At the end of the camp, students were asked to complete a questionnaire in order to compare the social and psychological outcomes displayed by the groups.

Results showed that the group of atypically dressed students rated themselves as less happy than their typical peers, reported a lower level of same color group identification (e.g., the atypical students reported a desire to switch color groups), but rated themselves as being more similar to their peers in the same color group than typical students rated themselves (Patterson & Bigler, 2007). These results indicated that atypical students were caught in a situation where they did not feel as though they fit in with the group that they are a part of; however, they tried even harder than their typical peers to find ways in which they are similar to those in the group to which they wished to belong. It is no surprise, then, that this level of psychic confusion and unease regarding their group affiliation would lead to the experiencing lower levels of happiness.

**Group Differences in Academic Achievement**

As discussed, behaving in a manner that is different from a student's peer group at large, including demonstrating a different level of academic achievement, can lead to negative outcomes including, but not limited to, bullying victimization. Based on these findings, it could be hypothesized that individuals who demonstrate above average academic interest and success in such a manner that separates that person from his or her social group may be at higher risk for bullying victimization.

Education research has consistently shown that an academic achievement gap exists between racial minority students and White students as well as between female students and male students in some subjects (Ladson-Billings, 2006). These differences have been reported and studied for nearly 50 years; suggesting that this gap has persisted long enough to become a cultural expectation. In an early example of this achievement gap, Coleman and colleagues (1966) reported that White students in American public schools were noticeably outperforming Black, Mexican American, Puerto Rican, Native American, and Asian students in 1$^{st}$ grade and in 12$^{th}$ grade on all measure academic abilities.

Countless interventions have been attempted in hopes of narrowing this achievement gap. Most interventions have targeted the level of support experienced by minority students who are failing academically in hopes of raising the minority groups' average achievement by reducing the number of students who are far behind their racial majority peers. Governmental initiatives like No Child Left Behind (NCLB) include an explicit goal of making sure that each child is at least minimally competent academically and forces schools to participate in standardized testing to help track students' progress towards the competency goal. This is not necessarily an inappropriate tactic. All students certainly do need to be able to complete basic academic tasks and these abilities are vital for future employment and promoting a fair and just society.

However, with the vast amount of resources used to target the improvement of these underperforming students, minority students who have achieved the basic academic skills may be overlooked. Plunker and colleagues (2010) investigated the educational progress of the highest achieving minority students compared to the progress of the highest achieving White students. These researchers examined students' reading and math achievement as measured by the National Assessment of Educational Progress (NAEP) scores taken during students' $4^{th}$ and $8^{th}$ grade years. These scores showed that that the gap between the highest achieving Black and Hispanic students and the highest achieving White students in both reading and math achievement had actually widened since the enactment of NCLB, even while the overall achievement gap between the races had begun to decrease.

The research team referred to this gap as the "excellence gap." This "excellence gap" has also been shown to exist for female students in comparison to male students in mathematics achievement when measured during both $4^{th}$ and $8^{th}$ grade (Plunker et al., 2010). For each of these groups of students, having their best students continue to lag behind the White students and male students, respectively, seems to suggest these groups will still experience some of the same occupational and leadership ceilings that these groups have experienced for years.

Despite increased awareness and attempted interventions to help reduce the gap, the differences in overall scholastic achievement between White students and minority students have largely remained (Barone, 2011; Good et al., 2008; Lee, 2002). Research still shows that, on average, Black and Hispanic students lag behind White and Asian students in every academic category, while girls perform noticeably lower in the areas of math and the physical sciences when compared to boys (Aronson et al., 1998). While there is much debate regarding the cause of the achievement gap, the fact that the gap has been maintained for such an extended period of time suggests that the academic performance levels consistent with the gap (e.g., White male students typically achieving higher than minority and female students) is the socially expected outcome. Perhaps some social stigma or pattern of interaction, like bullying, may be one factor contributing to the maintenance of the overall achievement gap.

## Math and Science Achievement

While the achievement gap between different groups' general academic functioning has garnered a large amount of interest, research, and intervention over the past 50 years, a more specific area of academic achievement has begun to receive more recent attention. Government and private agencies have begun tracking and advocating for increased educational standards and opportunities in math and science education. Frequently, this academic consideration has been broadened to also include technology and engineering education, which in association with math and science, has been dubbed the STEM fields. Broadly, STEM fields are defined as those areas of education pertaining to mathematics, the natural sciences, physical sciences, biological/agricultural sciences, engineering/engineering technologies, computer/information sciences, and the social sciences (Chen & Weko, 2009; United States Government Accountability Office, 2012).

This target focus on STEM education in general, and math and science instruction specifically, is the result of some troubling trends in the American education system and the future world job market. While America has traditionally prided itself on its education system as a way to prepare all young people to have the opportunity to achieve the American Dream, currently, compared to other 34 other industrialized nations, the United States ranks 25$^{th}$ in students' math achievement and 17$^{th}$ in science achievement (United States Department of Education, 2013). Furthermore, of the high school seniors who are interested in entering a STEM-related field, only 16% have been found to be proficient in math (United States Department of Education, 2013).

As schools have apparently slipped in their ability to prepare students in STEM-related areas, recent governmental and industrial examinations into the state of the American economy have shown that America is currently experiencing, or on the verge of experiencing, a shortage of workers capable of completing the math-, science-, and technology-based jobs that are now necessary to compete in the global economy (Office of Management and Budget, 2013). In the midst of this shortage, job and economic projections estimate that jobs in STEM-related fields in the United States will grow at a rate of 17% over 10 years compared to a growth rate of 9.8% for non-STEM fields (National Math & Science Initiative, 2013).

In order to help reverse this trend, in 2010 the U.S. government spent $3.1 billion funding 209 different programs that directly supported STEM education (United States Government Accountability Office, 2012). One of the most commonly referenced goals of this investment of public resources is to reduce the achievement, participation, and interest gaps in STEM fields between White students and racial minority students, as well as male and female students (United States Department of Education, 2013).

In general, White males have dominated math- and science-related training programs and careers for many years (Crowley, 1977; Sells, 1976). Employment and academic degree data published in the 1980s shows that, at least up to that point in American history, women and minorities significantly lagged behind White men in math and science participation. Data from 1986 indicated that, while women comprised 49% of the total American workforce, of the employed scientists, mathematicians, and engineers, only 15% were female. Even worse, Black and Hispanic workers each comprised only 2% of the number of employed scientists, mathematicians, and engineers (Oakes, 1990). Not only were White men found to be dominating the math- and science-related careers, they were also overrepresented in the math and science training programs based on the number of math- and science-related bachelor's degrees awarded by American colleges and universities. In 1985, Black students were awarded only 5% of the total scientific bachelor's degrees while Hispanic students were awarded only 3% of the total scientific bachelor's degrees (Oakes, 1990).

As the publication of these rates indicate, the gap between the races and genders in math and science participation and achievement has been noticed and studied for some time. But even with the exposure of the problem, data still suggest that the gap persists. Data from a nationally representative longitudinal data set show that White, Black, and Hispanic students all increased their level of math and science course taking from 1982 to 2004, suggesting, perhaps, that the awareness and resources paid to the issue of math and science education may be positively impacting the amount of student math and science participation. However, while enrollment rates for each of the races increased, the gap between the rates of enrollment between White and Black students actually widened, from an 8% difference in enrollment in 1982 to a 12% gap in 2004 (Dalton et al., 2007). This indicates that, while overall participation has increased, racial equality in math and science participation appears to be far from realized.

One way to track high school students' level of preparation for math- and science-related work is through the Advanced Placement (AP) exams. AP exams are nationwide tests administered to students typically after a year of intense high school instruction. The tests are scored on a scale to 1 to 5 where a score of 3 is considered passing. Many colleges and universities accept passing scores on an AP test as the equivalent of a college course.

Based on a report from The College Board, the group responsible for writing and distributing the Advanced Placement (AP) exams to schools around the nation, more girls are taking AP exams than ever before. In fact, girls took 55.6% of the tests administered in 2012 and 246,304 more girls participated AP testing than did boys (The College Board, 2013). But as this total participation has shifted in favor of female students, participation in math- and science-related classes, and subsequent rates of AP exam completion, remains remarkably in favor of male students. Out of the 34 total AP tests administered, 10 of those tests are related to math or science concepts (Biology, Calculus AB, Calculus BC, Chemistry, Computer Science, Environmental Science, Physics B, Physics C – Electricity and Magnetism, Physics C - Mechanics, and Statistics). Of the tests that are not considered math- or science-related, girls outnumber boys on 75%, or 18 out of 24, of the tests. However, on tests that are math- or science-related, girls outnumber boys on only 30%, or 3 out of 10, of the tests (The College Board, 2013). These participation rates suggest that, even while girls' participation in the most difficult high school classes outpaces boys overall, girls are still underrepresented in the math- and science-related classes to this day. Furthermore, since more girls are missing the early opportunities for training in math and science fields, they are less prepared than boys to enter math- and science-related majors in college and to enter the math- and science-based future work force.

Similarly, despite recent federal and state initiatives to improve minority student participation and success in math and science classes, minority students are still participating less often in math- and science-related AP exams. In 2013, White students accounted for roughly 54% of all students enrolled in United States high schools. White students, however, took nearly 69% of all the math- and science-related AP exams. Likewise, Asian students were also overrepresented in taking AP exams, as Asian students account for only 5% of the nation's total enrollment but took nearly 15% of the AP math- and science-related exams.

Black and Hispanic students, on the other hand, were underrepresented in the number of math- and science-related AP exams taken. While Black students account for roughly 16% of total enrollment, these students only attempted 5% of the math and science AP exams. Similarly, Hispanic students account for 22% of enrollment yet only attempted 6% of the AP math- and science-related exams (The College Board, 2013; United States Census Bureau, 2012).

What adds to the concern for Black and Hispanic students is that in 2012, Hispanic students passed only 41% of the AP exams they took, while Black students passed only 27% of AP exams they attempted (Simon, 2013). While this data is related to students' scores on all AP tests attempted and not only math- and science-related tests, if these rates of passage hold for the students' attempts on math- and science-related tests, only a very small fraction of Black and Hispanic students are leaving high school with the advanced education necessary to thrive in math and science fields later in life.

**Why the Gap in Math- and Science-Related Achievement?**

Like the achievement gap at large, the math and science achievement and participation gaps between minority students and female students have been a topic for debate for some time. Several different arguments have been made that attempt to explain why this gap exists and persists. Most theories assert that the gap is most likely caused by some problem located within the minority and female student, either an ability deficit based on inherent qualities or an inability to successfully tune out the societal stereotypes regarding their groups' math- and science-related abilities.

**Problem related to student deficiencies.** According to the inherent cognitive abilities hypothesis, minority and female students are born with lesser cognitive skills that prevent them from achieving at the same level as White male students in the areas of math and science. This notion that White males are inherently more capable than others is by no means a new idea (see Gould, 1981). The idea of generally deficient cognitive ability is frequently cited as a cause for the racial portion of the gap in overall academic achievement, and math- and science-related achievement by extension (Roth et al., 2001; Rushton & Jensen, 2005). To explain the gender-based differences in math and science achievement, research often turns to genetic differences in visuospatial abilities based on brain structure and chemistry (Deary et al., 2010; Spelke, 2005; Summers, 2005).

While there is still a large amount of debate surrounding these topics, recent research has begun to suggest that the inherent cognitive qualities of minorities and females are not different from those qualities presented by White males. Several of the landmark studies suggesting intrinsic gender-based differences in spatial skills have been shown to have significant methodological errors or reported results that have failed to be replicated (Spelke, 2005). Further damage to the intrinsic differences argument was provided by a review of 2nd through 11th grade standardized testing results from 10 states across the United States that showed that male and female students' scores on the math portion of the tests were statistically the same (Hyde et al., 2008). Similarly, minority students have been shown to demonstrate academic achievement levels on par with White students on a variety of academic measures when the influence thought to be caused by the internalization of societal norms and expectations regarding academic performance are minimized (Cohen et al., 1995).

Other hypotheses also consider the math and science achievement gap to be a problem that is caused primarily by minority and female students' reactions towards math and science tasks. One of the most influential hypotheses is the concept of stereotype threat. Stereotype threat is a subconscious process in which an individual is subtly made aware of his or her membership in a race- or gender-based group which society at large considers less capable in a given area (Good et al., 2008). Stereotype threat has been theorized to affect an individual in two ways. In a situation in which an individual is aware of a negative stereotype regarding a specific personality quality or ability level, he or she may experience stereotype threat when placed in a situation where the stereotype applies (e.g., a Black student taking an intelligence test). First, that individual will worry that his or her performance on the task will conform to the held stereotype about the group of which he or she is a member and, therefore, further the stigmatization of that group. However, stereotype threat also can bring about a more individually-based stress as the person attempting to overcome the held stereotype may fear that his or her performance on the task will confirm the individual's placement in the negatively stereotyped group (Wout, Danso, Jackson, & Spencer, 2008).

Steele and Aronoson (1995) initially studied this theory when they found that Black students performed worse on a standardized measure when they were told that the test was a measure on an individual's intellectual ability than they performed when told that the test was simply a laboratory problem-solving task. The researchers concluded that when Black students were made aware of the test's purpose, the students' stereotypes that Black students do poorly on measures of intellectual ability was activated. This finding suggests that those students who do not match the traditional expectations of the type of student who should be enrolled and/or be successful in math- or science-related classes may experience some psychological distress that could result in their underperforming in those classes.

**Problems related to the environment's reactions toward student.** The majority of the research regarding the math- and science-related achievement gap has focused on how the students who violate the math and science stereotypes relate to their environment and to the math and science tasks with which they are presented. Much less research has examined how the students' environment relates to them as they violate the traditional math and science stereotypes. However, while the research into how students who violate math and science expectations is limited, there is a strong literature base suggesting that students who stray from the academic norms and expectations can experience significant social repercussions.

One of the first studies to research these social repercussions found that students who violate the stereotypic racial norms for academic achievement and effort are often labeled as "acting White" by their peers (Fordham & Ogbu, 1986). By participating in stereotypically "White" behaviors, peers of minority students view the offending students as essentially rejecting their racial/ethnic heritage and, instead, place more importance on the traditional values of the dominant White culture. In their qualitative study that popularized the term "acting White" Fordham and Ogbu (1986) spoke to high ability yet typically underachieving students at an inner-city Washington D.C. high school. Students supplied a list of behaviors that were typically thought fit the description of "White" behaviors. The endorsed behaviors included "speaking standard English" and "getting good grades in school" (p. 11). Students were afraid that if they did achieve academically they would be labeled a "brainiac" and experience social stigmatization and isolation. In order to avoid this social isolation, which would certainly meet the qualifications of bullying, students often chose to actively underachieve and refuse to participate in academic activities (Fordham & Ogbu, 1986).

The "acting White" phenomenon was studied further in an affluent suburb at a racially integrated school in suburban Ohio (Ogbu, 2003). This school had a solid academic reputation, and the Black students who attended the school were typically upper-middle class. Still, Black students who attended the school reported feeling as though behaving in a way that would increase their ability to succeed academically (e.g. taking AP and Honors classes, paying attention and participating during lessons) would jeopardize their social standing and would often result in experiencing bullying from their peers (Ogbu, 2003).

Other studies that employed more quantitative-based methodologies have concluded that academic achievement, or lack of achievement, does appear to be related to students' level of popularity with their peers. Fryer (2006) investigated what he termed as the "pressure to be average." During this study, Fryer asked students to list the students in their class who they would consider their friend and also asked those students to describe their own academic grades. Fryer then determined which students were most frequently recorded as being someone's friend, and thereby determined each student's popularity. Fryer went on to compare each student's popularity rating and his or her reported level of academic achievement.

Results showed that for White students, increased academic achievement was associated with greater popularity, as the highest achieving students were generally rated as having the most friends. Results for minority students, however, differed noticeably. For Black students, increased academic achievement was slightly associated with increased popularity. Students achieving a grade point average (GPA) of 1.0 were rated as the least popular and ratings of popularity peaked for students who earned a GPA of 3.5. But when students achieved a GPA higher than 3.5, their reported level of popularity decreased as their GPA approached 4.0. Findings related to Hispanic students were even more startling. Again, increased GPA was somewhat associated with increased popularity, as students with GPAs of 1.0 were rated less popular than students with a

GPA of approximately 2.5. Interestingly, however, the reported popularity of a student plummeted when the student was found to achieve a GPA higher than 2.5. Popularity of these higher achieving Hispanic students fell to such a degree that students who achieve a 4.0 GPA were rated as having significantly fewer friends than even those students who earned a GPA of 1.0 (Fryer, 2006).

Whether described as “acting White” or the “pressure to be average,” academic achievement appears to be related to negative social outcomes for those racial minority students attempting to excel. Therefore, minority students appear to have an incentive to not pursue academic achievement or success if they value their social relationships and will often choose to underperform in academic areas, including math and science, in hopes of maintaining their social standing.

Female students have also been shown to experience negative social reactions when they violate the stereotypes of gender academic participation. In a qualitative study, Renold and Allen (2006) interviewed and gathered peer perspectives of three academically high achieving girls attending United Kingdom elementary schools. Two of the girls were shown to largely identify and conform to traditional gender norms and pressures. These girls discussed their attempts to appear both “bright” and “beautiful.” To maintain their conformity to social norms, these girls worked hard to maintain social relationships but were also submissive and self-deprecating as they attempted to minimize their academic achievements. The third girl, however, was described as much more masculine in the way that she presented herself and her academic success. This girl would admit to others that she was “clever” and took pride in the fact that she was a top student. In contrast to the other, more socially successful girls, this third student who did not match the traditional feminine stereotypes of humility and, instead, opted to openly discuss her success, was reported to be rejected by her peers and the authors noted several instances where she was clearly bullied (Renold & Allen, 2006).

While none of the studies presented discussed math or science achievement specifically, these studies illustrate the problems likely experienced by students who attempt advanced math- and science-related classes. Math and science education comes loaded with preconceived notions about who should, and should not, be able to achieve, and when students contradict those expectations and behave in an unexpected way, others often struggle to adjust to these violations and, instead, try to force the individual back to social conformity (Prentice & Carranza, 2002).

**Bullying of Females and Minorities Who Are Academically Different**

The behaviors that peers use to force those students who are violating racial and gender expectations inevitably fall under the definition of bullying. Again, there is a lack of research into the bullying experiences of students who specifically violate math- and science-based expectations. However, studies that examine the interaction of bullying and race in students who violate broad academics-related stereotypes can paint a clear picture about how students who violate math and science expectations are likely to be treated.

Research evidence has shown that minority individuals are significantly more likely to experience negative, "backlash effects" when they violate gender or racial stereotypes (Phelan & Rudman, 2010; Rudman & Fairchild, 2004). In two studies that manipulated the experience of stereotype compliance with the study participants, Phelan and Rudman (2010) and Rudman and Fairchild (2004) found that when individuals were placed in situations where they experienced a violation of racial or gender stereotypes, the individual who experienced the violation chose to punish the violator of the stereotype and attempted to sabotage that individual's chances for future success. These findings suggest that those students who violate the math- and science-based stereotypes may experience active resistance by others in their attempts to succeed in their math and science coursework.

While Phelan and Rudman (2010) and Rudman and Fairchild (2004) suggest that the resistance those students who violate traditional stereotypes may experience will be more subtle behaviors (e.g., sabotage, undermining abilities), other research has found that those who violate academic-related stereotypes are more likely to experience aggressive bullying victimization. Investigations using data gathered from a nationwide data set, the Educational Longitudinal Study of 2002 (ELS), found that Black students who displayed academic success were significantly more likely to experience bullying victimization than White students who demonstrated the same level academic success. The same trend was found with Hispanic students although the likelihood of bullying did not reach significance (Peguero & Williams, 2011; Williams & Peguero, 2013).

Using the same data set, researchers found that Black, Hispanic, and Asian students who participated in academic extracurricular activities (e.g., academic honor society, service organization) were significantly more likely to experience violent victimization than White students who participated in the same types of activities (Peguero et al., 2011). The authors of these studies claim that these findings are further evidence that Black students and, to some degree, Hispanic students are at a greater risk of bullying victimization when they display high rates academic achievement and higher rates of academic interest.

Even though this body of research does not mention math and science involvement specifically, it is clear that the academic qualities consistent with math and science interest, participation, and success are all associated with an increased risk of peer victimization for those students who violate the traditional expectations of students who should be participating in math and science education. Therefore, this literature supports the hypothesis that racial/ethnic minority students, especially Black and Hispanic students, as well as female students, should experience an increased rate of peer victimization if they choose to participate and succeed in math- and science-related education.

## Summary

School bullying is a significant problem, with between one-third and one-half of American students experiencing bullying victimization during their time at school (Bradshaw et al., 2007; Robers et al., 2012). Those students who experience bullying victimization have been shown to experience such negative outcomes as depression, loneliness, anxiety, serious suicide ideation, suicide attempts, and lower levels of self-esteem as compared to their non-bullied peers (Craig, 1998; Fekkes et al., 2006; Juvonen et al., 2000; Klomek et al., 2008; Kochenderfer & Ladd, 1996; Seals & Young, 2003). Despite a large amount of study and intervention, the problem persists throughout the nation. Bullying is typically defined as involving three distinct parts (Gottheil & Dubow, 2001). First, bullying is described as instrumental aggression and, therefore, serves to accomplish some goal of the aggressor (Espelage & Swearer, 2003). Second, the bullying episodes generally occur repeatedly over time (Monks & Smith, 2006). Third, a power differential must exist between the bully and his or her victim (Espelage & Swearer, 2003). Bullying has been shown to affect students of all ages (Crick et al., 1999; Kochenderfer & Ladd, 1996; Li, 2010; Lipinski & Crothers, 2014), genders (Archer & Coyne, 2005), ethnicities (Williams & Peguero, 2013), and levels of intelligence and academic achievement (Bishop et al., 2004).

Students especially at risk for experiencing bullying victimization are those children who are noticeably different from their peers (Patterson & Bigler, 2007). One area where some female and minority students demonstrate significant differences from the expectations held by their peers is when those students demonstrate a high level of academic achievement, especially in the areas of math and science. Traditionally, minority and female students have lagged far behind their White male peers in both achievement and participation in math and science courses, creating a consistent gap (Dalton et al., 2007; Oakes, 1990). This gap has become an increasingly important issue as American students have slipped in their math and science achievement compared to students in other developed nations (United States Department of Education, 2013) and as the global economy continues to demand more and more math and technology capable employees to sustain the workforce (Office of Management and Budget, 2013).

There is still a great deal of debate over what may cause the academic, and math and science by extension, achievement and participation gaps. Some theories claim that the gap is best explained by inherent differences in academic capabilities between minority groups, females, and White males (Roth et al., 2001; Rushton & Jensen, 2005). Others argue that the differences are largely attributable to minority and female students' fears that they will, in fact, fail at academic tasks and, therefore, confirm the academic stereotypes about their group and their own individual abilities, a phenomenon generally referred to as stereotype threat (Steel & Aronoson, 1995; Wout et al., 2008).

Still others believe that the problem these students experience is not based on their own perceptions of their academic ability, but rather is a result of others' reactions towards those students who have violated the academic expectations. Students who violate commonly held academic stereotypes frequently experience social stigmatization, often as a result of being perceived as "acting White" (Fordham & Ogbu, 1986; Ogbu, 2003), as well as both subtle and overt attempts to thwart the students' future ability to succeed academically and continue to violate the stereotype (Phelan & Rudman, 2010; Rudman & Fairchild, 2004). One of the most overt type of behaviors that students who violate stereotypes will experience is direct forms of bullying, which research has shown is more likely for Black and, to some extent, Hispanic students who achieve academic success (Peguero & Williams, 2011; Williams & Peguero, 2013).

The link between the experience of bullying, race, and students who achieve academically in untraditional areas is still unclear. Early findings of related studies suggest that those students who perform differently than most would expect, based on their racial and gender group affiliation, typically experience increased rates of bullying victimization. Peers' responses, specifically in regards to bullying, to differing levels of math and science achievement and participation demonstrated by minority and female individuals, however, have not been investigated. The current study attempts to fill this gap in the literature.

# CHAPTER 3 – Methods

## Research Questions

An investigation was completed in an attempt to answer the following research questions:

Research Question 1: Does a student's identification with a specific race impact his or her experience of peer bullying victimization after controlling for the student's ability, enjoyment, and confidence in mathematics?

> Hypothesis 1: After controlling for his or her mathematics ability, confidence, and enjoyment, a racial minority student will report higher frequencies of peer bullying victimization than a White student.

Research Question 2: Does a student's identification with a specific race impact his or her experience of peer bullying victimization after controlling for the student's ability, enjoyment, and confidence in science?

> Hypothesis 1: After controlling for his or her science ability, confidence, and enjoyment, a racial minority student will report higher frequencies of peer bullying victimization than a White student.

Research Question 3: Does a student's gender impact his or her experience of peer bullying victimization after controlling for the student's ability, enjoyment, and confidence in mathematics?

> Hypothesis 1: After controlling for her mathematics ability, confidence, and enjoyment, a female student will report higher frequencies of peer bullying victimization than a male student.

Research Question 4: Does a student's gender impact his or her experience of peer bullying victimization after controlling for the student's ability, enjoyment, and confidence in science?

Hypothesis 1: After controlling for her science ability, confidence, and enjoyment, a female student will report higher frequencies of peer bullying victimization than a male student.

**Data Source**

Data to answer the above questions were gathered from the Trends in International Mathematics and Science Study of 2011 (TIMSS 2011). The TIMSS surveys and assessments, which were first developed and administered in 1995, are now in their fifth iteration. The project is sponsored domestically by the United States Department of Education and the National Center for Education Statistics (NCES), and internationally by the International Association for the Evaluation of Educational Achievement (IEA). The TIMSS 2011 is comprised of a series of surveys administered to students, teachers, and administrators of participating schools as well as standardized math and science performance measures that gauge students' skills in the two academic subjects (Mullis, Martin, Ruddock, O'Sullivan, & Preuschoff, 2009).

The TIMSS 2011 was administered to fourth and eighth year (grade) students in 56 countries and education systems across the world, including schools in the United States. The TIMSS data is collected in order to help nations track students' math and science achievement, compare students' level of achievement around the world, notice trends in student capabilities or weaknesses, as well as illuminate different deficiencies with nations' school systems and curriculums (Mullis et al., 2009).

**Participants**

While the TIMSS 2011 was administered to students from countries around the world, for the purpose of this study, only the United States' sample will be used. Also, since rates of bullying behavior are often shown to peak in early adolescence (Nansel et al., 2001; Robers et al., 2012), the eighth grade sample of the TIMSS 2011 will be used for the present analyses. The following description of the selection of the TIMSS 2011 test sample is taken from the *U.S. TIMSS and PIRLS 2011 Technical Report and User's Guide* (Kastberg, Roey, Ferraro, Lamanski, & Erberber, 2013).

**Stratum Development.** The population of schools that would possibly be chosen to participate in the TIMSS 2011 was determined from two national databases collected by the NCES. The first database was the Common Core of Data (CCD), which listed all public schools in the United States that housed eighth grade students. The second database was the Private School Survey (PSS), which listed all private schools in the United States that housed eighth grade students. The information from these datasets was gathered by NCES during the 2007-2008 school year, which was the most current data available at the time of the construction of the TIMSS 2011. Information from these databases led to a total of 46,312 schools and 4,012,076 eighth grade students being considered eligible for participation in the TIMSS 2011.

From this total population eligible for participation, TIMSS 2011 designers then divided the schools into three distinct, explicit stratums. The first strata divided schools into four distinct groups based on their geographical location. Designers used the U.S. census geographical distinctions of Northeast, Midwest, South, and West to label the schools. The second strata divided schools into two groups depending on their poverty level.  Schools were labeled as “high poverty” if 50 percent or more of the school’s student population received free or reduced lunch, while schools with less than 50 percent of their student population receiving free or reduced lunch were labeled as “low poverty.” Finally, the third strata divided schools into two groups depending on their control status (e.g., public control vs. private control).

Next, within each of these developed explicit stratums, designers then implicitly stratified the population by three categorical stratification variables to further classify eligible schools.  The first variable chosen was the schools’ urban-centric locale and schools were classified as city, suburb, town, or rural.  The second variable chosen was the schools’ race/ethnic minority composition, with schools consisting of 15 percent or more of students described as Black, Hispanic, Asian and Pacific Islander, and/or American Indian and Native Alaskan being grouped together, while schools with fewer than 15 percent of minority students were being grouped together.  Finally, schools’ estimated eighth grade enrollment was used as a continuous variable.

**Sampling Procedure.** Based on these explicit and implicit strata, the final sample for the TIMSS 2011 was selected using a two-stage design.  The first stage utilized a stratified systematic sample of schools with sampling probabilities proportional to size (PPS), and the second stage involved selecting specific classrooms that would be sampled from selected schools.  During the first stage, the sampling design procedure allowed each of the schools eligible to participate in the study to have approximately the same probability of being selected to participate.  The TIMSS 2011 developers then attempted to assemble a sample of 600 schools that housed eighth grade students to participate in the study.

In accordance with the PPS method of sampling, target sample proportions based on actual population characteristics were calculated for each of the explicit and implicit stratum. In order to construct a sort ordered list of schools necessary to complete the stratified systematic sampling procedure, schools were first divided into groups according to each of their explicit stratums. Then, researchers calculated each school's "measure of size" (MOS) in order to create a self-weighting student sample. The MOS is proportional to each school's share of the target population based on overall student enrollment in eighth grade. Using the MOS, therefore, increased the probability that larger schools would be selected to participate in the study, which was both consistent with population characteristics and more cost effective for the researchers by collecting more student responses per school studied. Due to students who attended extremely small schools receiving very high weights and, therefore, inappropriately influencing the data, researchers determined that the minimum allowed MOS would be five.

Finally, within each explicit stratum, schools were rank ordered according to the two categorical implicit strata – locale and race/ethnic composition – as well as by the school's MOS in an alternating, or "serpentine," manner. This method of ranking schools ensured that schools adjacent to each other in the rank order were generally not substantially different on any of the implicit strata and never substantially different on more than one stratum.

In order to guard against potential school refusals to participate, substitute schools were also selected. For each school that was originally selected, two substitute schools were also identified. The first substitute school was the school immediately after the originally selected school in the rank order list, while the second substitute school was the school immediately before the originally selected school in the rank order list. This practice of determining substitute schools was considered appropriate given the high level of similarity between schools that are adjacent in the rank order.

After gaining consent from selected schools to participate in the TIMSS 2011 study, researchers then selected which classrooms would participate in the study. All eighth grade mathematics classes were listed as possible participants for the study. Schools were then asked if any of the math classes offered contained a majority of students who held a special classification (e.g., limited English proficiency; learning disability). Due to the TIMSS 2011 not providing any accommodations for students with special needs, these classrooms were removed from consideration of participation in the study. Next, if a classroom contained fewer than 15 students, that classroom was combined with other classrooms to create "pseudoclassrooms" of at least 20 students. If a selected school had two or more eighth grade mathematics classrooms, those classrooms were randomly assigned to participate in either the TIMSS 2011 or the co-occurring National Assessment of Educational Progress (NAEP) – TIMSS linking study. All students in the selected classrooms participated in the TIMSS 2011.

**Final Sample.** Completion of the TIMSS 2011 assessments and questionnaires was completed between April 2011 and June 2011. Of the 600 schools contacted to participate in the TIMSS 2011 study, 574 were rated as eligible to participate. Of those schools, 499 original and 2 substitute schools agreed to participate, resulting in 501 total schools participating; an 87% participation rate. In those 501 participating schools, there were a total of 5,199 mathematics classrooms, 4,663 of which were rated as eligible to participate. Of this eligible total of classrooms, 557 were asked to participate and all agreed, resulting in a 100% participation rate. Finally, the participating classrooms held a total of 11,860 students. Of these students, 11,160 students were eligible to participate. On the day of administration, 687 students were absent, resulting in a total of 10,480 students being assessed; a 94% participation rate. The characteristics of this final sample was extremely similar to the population characteristics on each of the explicit and implicit strata as none of the sample proportions were more than 0.3% different from population characteristics.

## Instrumentation

As previously stated, the TIMSS 2011 consists of a variety of measures including: a student questionnaire, a teacher questionnaire, a school questionnaire, a questionnaire assessing schools' curricula, a mathematics achievement assessment, and a science achievement assessment. For the purpose of this study, information gathered from the student questionnaire and both the math and science achievement measures will be used for analysis.

### Context Questionnaires

While the goal of the TIMSS 2011 is to measure and track student performances in math and science academic work, creators of the TIMSS realize that student academic achievement does not occur in a vacuum. Research and learning theories have shown for many years that student academic performance is impacted by countless intrapersonal, interpersonal, and contextual factors (e.g., Bandura, 1977; Bronfenbrenner, 1977). In order to assess some of these factors, the TIMSS 2011 creators developed contextual questionnaires in order to gather information regarding four broad domains: national and community contexts, school contexts, classroom contexts, and student characteristics and attitudes. Items pertaining to these four domains are dispersed across the different contextual questionnaires in order to gain the perspectives and insights from multiple involved sources (Mullis et al., 2009).

These questionnaires were developed based on the TIMSS National Research Coordinators' (NRCs) as well as the TIMSS 2011 Questionnaire Item Review Committee (QIRC) members' interpretations of the most critical contextual factors that impact students' academic performance according to relevant educational literature (Mullis et al., 2009). The process of constructing these questionnaires began in February 2008 and continued until August 2010, when the final versions of the questionnaires were approved and distributed to participating schools (Mullis & Martin, 2011).

Student context questionnaires were intended to be completed with paper and pencil, and items regarding student background information provided participants with various options to endorse (e.g., self-report of gender, race/ethnicity, family income). Other items were written as a part of a developed scale that were intended to provide insights into a specific construct that TIMSS 2011 developers considered important to students' ability to learn and perform in math and science. These items were written as statements to which participants could respond by endorsing one of four descriptive anchors. These anchors were "Disagree a lot," "Disagree a little," "Agree a little," and "Agree a lot" (Martin, Mullis, Foy, & Arora, 2011).

Prior to final approval, the contextual questionnaires were field tested in order to assess their appropriateness. Sites were chosen to participate in the field tests of the scales in the same method that the general sample was chosen, and resulted in roughly 30 schools and 200 students being field tested from each participating nation. Field testing was completed in order to determine the unidimensionality, the internal consistency, and the validity of the measured scales, as well as the scales' relationship to student achievement (Mullis & Martin, 2011).

Unidimensionality of the scales was deemed an important quality of the context questionnaires due to TIMSS 2011 developers' goals of using the 1-Parameter IRT measurement model during the analysis of the scales' results. One requirement of this measurement model is that all items in a developed scale must be related to a single underlying construct. In order to measure a scale's unidimensionality, the field test data from all participating students was aggregated and a Principle Components Analysis was completed. If a scale was found to have multiple factor loadings, the scale was revised to help align the items into a single factor structure, or the scale was eliminated. Similarly, if an item did not show a factor loading of 0.3 or higher on the principle underlying construct, that item was eliminated from the scale.

Internal consistency, or reliability, of the scales was also calculated after the field testing procedure. TIMSS 2011 developers established a Cronbach's Alpha coefficient of 0.7 as the cutoff for appropriate reliability. Scales whose coefficient did not reach the 0.7 limit were considered for elimination from the questionnaire.

Finally, TIMSS 2011 developers also examined the different scales' relationship with measured student achievement. As stated above, TIMSS 2011 NRCs and QIRC members added questions and scales to the student questionnaire based on the topics' research-based support for positive relationships with academic success. Therefore, the question scales developed should show a positive relationship with student achievement on the math and science measures, which would also provide further support for the validity of the scales.

To measure this relationship, developers first coded student responses to a numerical scale that allowed the scale to have an approximate mean of 10 with a standard deviation of 2. For example, on a scale with five items, the different score anchors would be recoded to a scale from 1 to 4, with “Disagree a lot” being coded as 1 and “Agree a lot” being coded as 4. This would lead to a maximum possible of 20 and a minimum possible score of 5. All scales were coded in such a way that led to higher scores being related to a more supportive learning environment.

A distribution of scores was then created based on the results gathered during field testing. Using this distribution, developers created three groups of scores; scores falling in the top 25% of responses, those scores falling in the middle 50% of scores, and scores falling in the bottom 25% of scores. Each student’s overall average score on the TIMSS 2011 math and science assessments was then calculated and plotted against these contextual scales that rated the supportiveness of the student’s environment. Based on the developers’ theory, it was expected that students reporting higher levels of support would score higher on the achievement measures.

The large majority of developed scales did show the expected relationship between student achievement, and their responses on the contextual questionnaires and were, therefore, considered valid scales. Some scales, however, did not display the desired positive relationship and were eliminated from the questionnaire. After the results of field testing, the TIMSS 2011 NRC approved questionnaires containing a total of 33 scales with nine scales found on the student questionnaire, eight scales on the school questionnaire, nine scales on the questionnaire completed by math teachers, and nine scales on the questionnaire completed by science teachers (Mullins & Martin, 2012).

**Study relevant scales.** Not all scales included in the context questionnaires will be used in this study and, therefore, will not be discussed further. The following five scales will be used during the subsequent analyses.

***Student Bullied at School Scale.*** The TIMSS 2011 “Student Bullied at School Scale” consists of six items and includes statements regarding whether students have experienced different types of bullying behaviors. Students are asked to respond to each statement by endorsing an anchor regarding their frequency of experiencing the bullying behavior. The anchors provided are “Never,” “A few times a year,” “Once or twice a month,” and “At least once a week.”

The items used on this bullying scale are clearly related to one of the bullying behavior constructs that has been discussed above. The first item, “I was made fun of or called names” is related to verbal bullying (Camodeca & Goosens, 2005). The second item, “I was left out of games or activities by other students” is related to social aggression (Archer & Coyne, 2005). The third item, “Someone spread lies about me” is related to relational aggression (Cullerton-Sen & Crick, 2005). The fourth item, “Something was stolen from me” is related to physical aggression (Card et al., 2008). The fifth item, “I was hit or hurt by other student(s); (e.g., shoving, hitting, kicking)” is also related to physical aggression (Card et al., 2008). Finally, the sixth item, “I was made to do things I didn’t want to do by other students” is related to social aggression (Archer & Coyne, 2005).

Based on the responses from students in the United States sample, this scale was found to have appropriate reliability ($\alpha = .78$) and the scale explained roughly 48% of the variance in student achievement. The individual items of the scale also demonstrated moderate to strong factor loadings, as items 1-6 showed correlations of $r = .75$, $r = .71$, $r = .72$, $r = .63$, $r = .72$, and $r = .62$, respectively. This scale also showed a small, positive correlation with students' mathematics achievement ($r = .04$) but no discernable relationship with students' science achievement ($r = .00$; Mullins & Martin, 2011).

***Confidence and enjoyment in math and science scales***. The TIMSS 2011 student context questionnaire contains scales that assess students' opinions and perceptions of their own skills and level of enjoyment in math and science. Each of these scales consist of a series of statements, and students are asked to endorse one of the four qualitative anchors, "Agree a lot," "Agree a little," "Disagree a little," and "Disagree a lot." The first scale is titled "Student Likes Learning Mathematics Scale" and contains five statements that attempt to assess how much students enjoy the subject of mathematics (e.g., "I enjoy learning mathematics," "I learn many interesting things in mathematics"; Table 1). The second scale is titled "Student Confident in Mathematics Scale" and contains nine statements that attempt to assess students' level of confidence in mathematics (e.g., "I usually do well in mathematics," I learn things quickly in mathematics"; Table 1). Students were then asked to respond to the exact same items a second time, however, responding with their opinions and perspectives about their abilities and enjoyment in science courses. The titles of these scales are "Student Likes Learning Science Scale" (Table 1) and "Student Confident in Science Scale" (Table 1), respectively.  Each of these four scales demonstrate strong reliability and validity, as well as moderate to strong factor loading coefficients (Table 1; Mullins & Martin, 2011).

**Mathematics Assessment**

The developers of the TIMSS 2011 view mathematics education to be extremely important in helping individuals becoming effective citizens and successful in the workplace.  To these developers, math education involves not only knowing math facts but also the ability to reason mathematically and apply mathematics skills to solve problems in everyday life.  In order to assess both students' math facts and math thinking skills, the TIMSS 2011 Mathematics Assessment is divided into two broad categories; the Content Domains and the Cognitive Domains.

**Domains**.  For eighth grade students, the TIMSS 2011 consists of four content domains.  The first content domain is Number.  The Number domain accounts for 30% of the questions on the mathematics assessment.  According to the description of the assessment frameworks prepared by Mullis and colleagues (2009), "the number content domain includes understanding of numbers, ways of representing numbers, relationships among numbers, and number systems" (p. 30).  The second content domain is Algebra. The Algebra content domain accounts for another 30% of the questions on the mathematics assessment.  This content domain is described as including items that require "recognizing and extending patterns, using algebraic symbols to represent mathematical situations, and developing fluency in producing equivalent expressions and solving linear equations" (Mullis et al., 2009; p. 32).

The third content domain is Geometry.  The Geometry domain accounts for 20% of the questions on the mathematics assessment.  Questions involving this content

Table 1

*Reliability, Factor Loading, and Achie[illegible]ics of Liking [illegible]ce Scales*

| Scale | Cronbach's Alpha Reliability Coefficient | % of Variance Explain | Factor Load Item 1 | Factor Load Item 2 | Factor Load Item 3 | Factor Load Item 4 | Factor Load Item 5 | Factor Load Item 6 | Factor Load Item 7 | Factor Load Item 8 | Factor Load Item 9 | Correlation with Subject Achieve |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Student Likes Learning Math Scale** | 0.89 | 69 | 0.89 | 0.78 | 0.81 | 0.77 | 0.90 | -- | -- | -- | -- | 0.23 |
| **Student Confident in Math Scale** | 0.89 | 55 | 0.77 | 0.74 | 0.81 | 0.81 | 0.71 | 0.76 | 0.61 | 0.61 | 0.79 | 0.43 |
| **Student Likes Learning Science Scale** | 0.88 | 68 | 0.88 | 0.75 | 0.81 | 0.79 | 0.90 | -- | -- | -- | -- | 0.27 |
| **Student Confident in Science Scale** | 0.89 | 53 | 0.77 | 0.72 | 0.76 | 0.80 | 0.68 | 0.75 | 0.69 | 0.67 | 0.71 | 0.35 |

domain require students to “analyze the properties and characteristics of a variety of two and three-dimensional geometric figures, including lengths of sides and sizes of angles, and to provide explanations based on geometric relationships” (Mullis et al., 2009; p. 34). Finally, the fourth content domain is Data and Chance. The Data and Chance domain accounts for the final 20% of questions on the mathematics assessment. Developers describe this domain as asking students to “organize data that have been collected by oneself or others and how to display data in graphs and charts that will be useful in answering questions that prompted the data collection” (Mullis et al., 2009; p. 36).

While these content domains describe the mathematical topic each question assesses, the cognitive domains illustrate the level and style of thinking that students are required to demonstrate in order to solve a question. TIMSS 2011 developers describe three cognitive domains. The first cognitive domain is Knowing and accounts for roughly 35% of the items asked. This domain “covers the facts, concepts, and procedures students need to know” (Mullis et al., 2009; p. 40). Questions in this domain ask students to demonstrate their ability to recall important definitions and number properties, recognize mathematical information, carry out algebraic computations, retrieve information from graphs and tables, use measuring instruments with appropriate units, and classify objects, shapes, and numbers.

The second cognitive domain is Applying and accounts for 40% of the items asked. This domain “focuses on the ability of students to apply knowledge and conceptual understanding to solve problems or answer questions” (Mullis et al., 2009; p. 40). Items that target this domain force students to select an appropriate method for solving a problem, display mathematical information in a chart or graph, construct an

Table 2

*Summary Statistics and Standard Errors for Proficiency on the Mathematics Content and Cognitive Scales*

| **Scale** | Sample Size | Mean Proficiency | Standard Deviation | Jackknife Sampling Error | Overall Standard Error |
|---|---|---|---|---|---|
| **Overall** | 10,480 | 509.485 | 77.047 | 2.575 | 2.633 |
| **Number** | 10,480 | 513.825 | 84.907 | 2.781 | 2.991 |
| **Algebra** | 10,480 | 511.827 | 74.392 | 2.533 | 2.576 |
| **Geometry** | 10,480 | 484.777 | 81.554 | 2.694 | 2.722 |
| **Data & Chance** | 10,480 | 527.263 | 103.074 | 3.148 | 3.273 |
| **Knowing** | 10,480 | 519.300 | 78.522 | 2.635 | 2.690 |
| **Applying** | 10,480 | 503.057 | 81.467 | 2.682 | 2.809 |
| **Reasoning** | 10,480 | 503.408 | 82.451 | 2.689 | 2.719 |

(Martin & Mullis, 2012)

appropriate equation or diagram to solve a problem, implement a set of mathematical instructions, and to solve math problems that they would likely experience in a math class.

The third cognitive domain is Reasoning and accounts for 25% of the items asked. This domain "goes beyond the solution of routine problems to encompass unfamiliar situations, complex contexts, and multistep problems" (Mullis et al., 2009; p. 40). Items included in the reasoning domain require students to analyze relationships between variables and make appropriate inferences from given information, restate results of a mathematical procedure in general, more understandable terms, integrate and synthesize different pieces of information and mathematical ideas, justify their own mathematical procedures, and solve novel, non-routine problems.

These content and cognitive domains were considered as a scale after the data was gathered. Information regarding each scales' sample size, students' mean proficiency, standard deviation, sampling error, and overall standard error is provided in Table 2. The overall mean of the mathematics assessment is 500, with an overall standard deviation of 100. By tracking students' responses through both these content and cognitive driven domains, the TIMSS 2011 allows researchers to determine not only the specific math skills that students are being taught, but also the students' ability to think mathematically and apply the skills that they have. This information can lead to a deeper understanding of the strengths and weaknesses in a given nation's developed curriculum, or in the implementation of that curriculum, and can help in the development of targeted suggestions on how to fix any problems that are discovered.

**Science Assessment**

Similar to their stance on mathematics, TIMSS 2011 developers also view a solid base of science knowledge and understanding to be vital in preparing students to live productive, informed lives. Scientific understanding, as well, necessitates that students know basic facts but are also prepared with the cognitive skills to process through scientific information and apply the information that they have gathered to solve problems. Mirroring the structure of the mathematics assessment, the science section of the TIMSS 2011 also contains four Content domains pertaining to different areas of science as well as the same three Cognitive domains that were included in the mathematics section.

**Domains.** The first content domain in the science assessment is Biology and questions pertaining to this domain account for 35% of questions asked. In this domain, students are expected to answer questions from six general topics: (1) characteristics, classification, and life processes of organisms, (2) cells and their functions, (3) life cycles, reproduction, and heredity, (4) diversity, adaptation, and national selection, (5) ecosystems, and (6) human health (Mullis et al., 2009; p. 64). The second domain in the science assessment is Chemistry and questions about this domain account for 20% of the items administered. In this domain, students are expected to answer questions regarding three general topics: (1) classification and composition of matter, (2) properties of matter, and (3) chemical change (Mullis et al., 2009; p. 69).

The third content domain in the science assessment in Physics and questions pertaining to this domain account for 25% of items administered. In this domain, students are asked about five broad topics: (1) physical state and changes in matter, (2) energy transformations, heat, and temperature, (3) light and sounds, (4) electricity and magnetism, and (5) forces and motion (Mullis et al., 2009; p. 72). Finally, the last content domain in the science assessment asks questions related to Earth Science, and questions regarding this topic account for 20% of items assessed. The TIMSS 2011 developers note that, while there is no single curriculum that is consistent across nations when it comes to Earth Science, the topics that the developers chose to include do appear to be held universally. Furthermore, the information assessed is frequently taught to students by their eighth grade year. There are four topics assessed in this domain including: (1) Earth's structure and physical features, (2) Earth's processes, cycles, and history, (3) Earth's resources, their use and conservation, and (4) Earth in the solar system and the universe (Mullis et al., 2009, p. 76).

While the science content domains measure students' knowledge of scientific facts and processes, the science assessment cognitive domains measure students' abilities to process through scientific thinking and use their scientific knowledge in new situations. The first cognitive domain on the science assessment is Knowing, and this domain accounts for 35% of the items administered. On questions related to this domain, students are asked to demonstrate their abilities to recall or recognize science facts regarding relationships, structures, and processes, to define or identify scientific terms, describe organisms, physical materials and science processes, to support or clarify statements of facts or concepts with appropriate examples, and to demonstrate their knowledge of how to use scientific instruments.

The second cognitive domain on the science assessment is Applying, and this domain accounts for 35% of items administered. This domain assesses students' abilities to apply their scientific knowledge in straightforward situations. Students are asked to identify similarities and differences between organisms, use models to aid in the demonstration of a scientific concept, relate knowledge of an underlying concept to an observed phenomenon, interpret presented information in light of scientific concepts, find solutions when presented with a direct application of concept, and explain an observation or natural phenomenon.

Finally, the last cognitive domain in the science assessment is Reasoning, and items assessing this domain account for 30% of items administered. This domain measures students' abilities to analyze problems, integrate or synthesize scientific information, hypothesize or predict outcomes based on scientific experience or observation, design appropriate investigations, draw conclusions from gathered data, generalize conclusions beyond a given situation, evaluate the advantages and disadvantages of different scientific processes or decisions, and justify a conclusion using evidence and scientific understanding (Mullis et al., 2009).

Again, just like the mathematics assessment, the content and cognitive domains in the science assessment will be combined and calculated as a scale that can be used to track student progress and compare results across time and nationalities. Information regarding each scales' sample size, students' mean proficiency, standard deviation, sampling error, and overall standard error is provided in Table 3. The science assessment also has an overall mean of 500 with an overall standard deviation of 100.

**Assessment Construction**

The TIMSS NRC developed the frameworks for the mathematics and science assessments before each specific assessment was constructed. After the frameworks had been agreed upon, TIMSS developers contacted small groups of individuals to write individual items for the upcoming TIMSS 2011 assessments. These individuals were instructed to write items pertaining to a specific content and cognitive domain and were also given explicit guidelines to follow as they wrote the new items. Most notably, participating question writers were reminded to avoid nationality-, race/ethnicity-, and gender-based bias as they composed their questions. Item authors were also asked to ensure that the language used in the question text was simple and direct enough to be easily translated into a wide variety of languages. TIMSS developers were planning on adding approximately 85 new items to both the math and science assessments for the TIMSS 2011 administration (Mullis & Martin, 2011).

Questions on the TIMSS 2011 were written in one of two formats. The first format was multiple-choice. These items required a student to identify a correct response to a given situation out of four provided answer options. Item developers were instructed to compose direct questions with no extraneous information and provide three plausible

Table 3

*Summary Statistics and Standard Errors for Proficiency on the Science Content and Cognitive Scales*

| **Scale** | Sample Size | Mean Proficiency | Standard Deviation | Jackknife Sampling Error | Overall Standard Error |
|---|---|---|---|---|---|
| **Overall** | 10,480 | 524.602 | 81.401 | 2.485 | 2.552 |
| **Biology** | 10,480 | 530.247 | 85.509 | 2.498 | 2.538 |
| **Chemistry** | 10,480 | 519.753 | 90.632 | 2.616 | 2.639 |
| **Physics** | 10,480 | 513.311 | 80.482 | 2.388 | 2.499 |
| **Earth Science** | 10,480 | 533.133 | 92.494 | 2.689 | 2.771 |
| **Knowing** | 10,480 | 527.117 | 91.451 | 2.621 | 2.842 |
| **Applying** | 10,480 | 464.245 | 96.201 | 2.104 | 2.112 |
| **Reasoning** | 10,480 | 523.547 | 80.750 | 2.390 | 2.469 |

(Martin & Mullis, 2012)

distracters along with the one correct, best answer. Multiple-choice items were all worth one point in the scoring of the assessment and students should be expected to complete the item in one to three minutes.

The second item format is constructed-response. Constructed-response items were typically items that required students to provide some numerical response without the aid of provided answer options. Again, item developers were instructed to write clear, direct questions with easily accessible language and subject matter. Along with the item text, developers were required to write a scoring guideline for any constructed-response item that would allow for clear and consistent scoring of the item. Constructed-response items could be worth either one or two total points. If the item was worth two points, the student would also be able to receive one point in partial credit which would also be discussed in the scoring guidelines. Constructed-response items were also intended to be completed in one to three minutes (Mullis & Martin, 2011).

After the item developers had submitted their written questions, the new items were field tested to assess each item's appropriateness. The field testing procedure included 178 new mathematics items and 176 new science items. The field testing assessments were administered to 60,376 students from 44 different countries. This allowed TIMSS 2011 developers to determine which of the approximately 85 new math and 85 new science items best assessed students' skills in each of the content and cognitive domains (Martin & Mullis, 2011).

After field testing, TIMSS 2011 NRCs could begin to develop the final assessment booklets that would be administered to students. A significant issue, however, was that not only did roughly 170 new items need to be included in the TIMSS 2011 assessment, but, given the TIMSS stated goal of tracking education trends, a significant number of items from the previous TIMSS 2007 assessment needed to be administered in this iteration as well in order to accurately track student achievement over time. Furthermore, based on plans to analyze TIMSS 2011 results with IRT-based analyses and report in depth current and past trends, a very large number of assessment items were required to meet the assumptions of the various tests and achieve the necessary statistical power to make the desired inferences (Kastberg et al., 2013). Therefore, after field testing and determining which new items should be added to the assessment, including the items saved from the TIMSS 2007 assessment, the final TIMSS 2011 assessment consisted of 91 new mathematics items, 126 mathematics trend items from TIMSS 2007, 92 new science items, and 125 science trend items from TIMSS 2007 for a total of 434 assessment items across both subjects (Mullis & Martin, 2011).

With such a large number of assessment items, if each student was expected to complete each item at even one minute per item, the TIMSS 2011 assessment would take over seven hours to complete, which was considered unfeasible. To combat this issue, the TIMSS 2011 developers used a matrix-sampling approach and divided the total assessment into 28 blocks of items, 14 blocks containing mathematics questions and 14 blocks containing science items. Each block contained 12 to 18 items and could be completed in 22.5 minutes on average. Of the 28 blocks, 16, eight in each subject, contained trend items from the TIMSS 2007 assessment leaving 12 blocks, six in each subject, to hold the newly developed items. These blocks were constructed in such a way that each block contained roughly the same proportion of content and cognitive domains as found in the overall item pool (Mullis et al., 2009).

From the 28 blocks of assessment items, TIMSS 2011 developers created 14 booklets of items that would be administered to students. Each booklet consisted of four blocks of items, two blocks of mathematics items and two blocks of science items. Each item block was included in two separate booklets to allow for linking student responses across booklets. All booklets, except for one, contained one item block of newly developed questions for both math and science and one TIMSS 2007 trend item block for both math and science. One booklet was composed entirely of trend items from TIMSS 2007 (Mullis et al., 2009).

In half of the developed booklets, the two mathematics blocks were provided first while in the other half of booklets, the two science blocks were provided first. Booklets were dispersed to students in a sampled classroom in a predetermined order to ensure that the overall student ability levels were equivalent across booklets. During an assessment session, students were given 45 minutes to complete the first section of the assessment (e.g., two math or science blocks depending on the booklet) and then allowed to take a break. Following the break, students were given an additional 45 minutes to complete the second portion of the assessment and then allowed a second break before having 30 minutes to complete the student context questionnaire (Mullis et al., 2009).

**Plausible Values.** While the process of creating item blocks allowed for a more reasonable amount of time to complete the assessment, it severely limited the number of items that each student completed. By limiting the number of items that each student completed, developers realized that they would not have enough data to accurately calculate desired scale scores involving the content and cognitive domains. In order to generate enough data, TIMSS developers decided to conduct plausible-values methodology to supplement the observed data gathered from each student.

Plausible values are estimates of what a student would have scored on all the items in the assessment had he or she taken the entire assessment. This estimate is calculated for each student individually based on two general sources of information; a student's actual responses to administered items and that student's background information. The individual student is then matched with other students who scored in a similar manner on other items in the assessment and reported similar background information. Based on the responses of the matched students on the items that the original student had not attempted, an empirically derived distribution of possible responses on those items is created.

Because the estimation is based on a distribution of possible responses, however, a single estimation point would likely lead to skewed results. Therefore, five random draws of possible values for each item were taken. Any analyses using this data must, then, be run five times, each time using different results from these random draws and the final results of the five analyses should be averaged to create the most accurate estimate of the student's scale score (Kastberg et al., 2013).

## Research Design

### Operational Definition of Variables

For the purpose of this study, the following operational definitions for the variables used were developed. The dependent variable of the study will be referred to as Peer Bullying Victimization. Peer Bullying Victimization is defined as any peer interaction in which a more powerful individual does harm to another individual over an extended period of time with the intent of accomplishing some goal (Gottheil & Dubow, 2001). The Peer Bullying Victimization variable will be measured as students' responses to the six-item TIMSS 2011 scale called the Students Bullied at School Scale that was described above. Students were asked to respond to six statements related to bullying victimization. In response to these statements, students were instructed to indicate the frequency to which these statements apply to their own lives. Students were given the option of endorsing either "Never," "A few times a year," "Once or twice a month," or "At least once a week." This variable will provide insight into students' overall reports of bullying victimization.

The independent variables used in this study will be defined as follows. The first independent variable will be referred to as Gender. This variable is defined as a student's self-report of his or her dichotomous male/female gender. The information for this variable will be gathered from a question on the Student Context Questionnaire of the

TIMSS 2011, which asks students, “Are you a girl or a boy?” and then provides students with ovals to fill in to indicate their gender as either “Boy” or “Girl.”

The second independent variable will be referred to as Race/Ethnicity. This variable is defined as a student’s self-report of his or her race/ethnicity. The information for this variable will be gathered based on student responses on two questions from the TIMSS 2011 Student Context Questionnaire. The first relevant question asked students, “Are you Hispanic or Latino?” and instructed students to fill in one oval indicating either, “Yes, I am Hispanic or Latino” or “No, I am not Hispanic or Latino.” The next item on the questionnaire asked students, “Which of the following best describes you?” and gave students the option of filling in one or more of the ovals indicating that the student identified as “White,” “Black or African American,” “Asian,” “American Indian or Alaska Native,” and/or “Native Hawaiian or other Pacific Islander.”

The first and second variables that will be used as covariates in the following analyses will be referred to as Enjoyment of Mathematics [Science]. For this study, enjoyment of the specific subject is defined as the level to which a student reports liking or finding the subject interesting. Information for this variable will be gathered from the two, five-item TIMSS 2011 scales called Students Like Learning Mathematics [Science] Scale that are described above. Students were asked to respond by indicating their level of agreement with five statements regarding their level of enjoyment in mathematics or science. They were given the option of endorsing either “Agree a lot,” “Agree a little,” “Disagree a little,” or “Disagree a lot.”

The third and four covariates will be referred to as Confidence in Mathematics [Science]. For the purposes of this study, student confidence is defined as a student’s belief that he or she can understand and succeed in his or her mathematics or science coursework. Information for this variable will be gathered from the two, nine-item TIMSS 2011 scales called Students Confident in Mathematics [Science] Scale that are described above. Students were asked to respond to a variety of statements regarding their own perceptions of their performance as well as how others view their performance in mathematics and science courses. Participants were instructed to respond to these statements by indicating to what extent they agree with each statement by endorsing either “Agree a lot,” “Agree a little,” “Disagree a little,” or “Disagree a lot.”

The fifth covariate will be referred to as Mathematics Ability. In this study, a student’s mathematics ability will be defined as the rate at which a student is able to answer novel age and grade appropriate mathematics questions on a standardized mathematics assessment. Information for this variable will be gathered from the TIMSS 2011 mathematics assessment. The TIMSS 2011 mathematics assessment is further delineated into different domains that measure specific areas and concepts of mathematics. However, because there is no theoretical basis for one domain of mathematics being a better predictor of bullying victimization than another domain, only each student’s overall mathematics score will be used for analysis.

Finally, the sixth covariate used in this study will be referred to as Science Ability. A student’s science ability will be defined as the rate at which a student is able to answer novel age and grade appropriate science questions on a standardized science assessment. Information for this variable will be gathered from the TIMSS 2011 science assessment. The TIMSS 2011 science assessment is further compartmentalized into different domains that measure specific areas and concepts of science. However, because there is no theoretical basis for one domain of science being a better predictor of bullying victimization than another domain, only each student's overall science score will be used for analyses.

**Type of Analysis**

In order to determine if racial minority students' participation in STEM fields is related to an increased report of bullying victimization, a quantitative analysis of the results of the TIMSS 2011 survey will be completed. However, because the surveys and assessments administered by the TIMSS 2011 were not constructed specifically for this research project or to answer the current research questions, the analyses completed in this study should be considered secondary analyses. A research endeavor is considered secondary analysis if the data analyzed to answer the research question was not directly collected to answer those specific questions (Smith, 2008). This includes research that re-analyzes previously collected data using new, often more sophisticated statistical techniques to answer the originally designed research questions, as well as research that uses previously collected data to answer newly developed research questions (Smith, 2008).

Secondary analysis has been used as a research tool for some time and the results gathered through this technique have led to important discoveries in education and social science research (Glass, 1976). While secondary analysis techniques do contain certain limitations that must be considered, researchers still argue that the process of analyzing previously collected data can be an appropriate and informative way to determine patterns and draw conclusions regarding education, generally, and when working with academically interested and gifted students, specifically (Mueller & Hart, 2010).

**Construct Validity**

Due to the use of secondary analysis in this procedure, only the information that was gathered by the TIMSS 2011 surveys is available to be analyzed. With the presence of this limitation in the available data, it is important to investigate the construct validity of the procedure. For a research study to demonstrate good construct validity, all independent and dependent variables measured in the study should accurately reflect the theoretical constructs described in the relevant literature (Wampold, 2006).

The construct validity of the current study must be considered since the current researcher had no involvement in the development of the TIMSS 2011 surveys and the gathered data may or may not accurately match the current researcher's constructs that are intended to be measured (Kluwin & Morris, 2006). Typically, in order to help ensure accurate and valid research practices, researchers are instructed to develop specific research questions prior to collecting data in order to ensure that the data gathered can answer the questions asked (Booth, Colomb, & Williams, 2008). Since the research questions for this investigation were developed after the collection of the data, it is important to carefully consider whether or not the information available in the dataset is applicable and able to accurately answer the current questions.

First, the construct validity for the dependent variable of Peer Bullying Victimization is assessed. As discussed earlier, the questions posed to students about their experiences of bullying victimization are directly related to the literature-based definitions of different bullying behaviors including verbal aggression (Camodeca & Goosens, 2005), social aggression (Archer & Coyne, 2005), relational aggression (Cullerton-Sen & Crick, 2005), indirect aggression (Richardson & Green, 2006), and physical aggression (Card et al., 2008). Due to the strong construct validity of these individual question items, students' responses should be considered to be an accurate reflection of their actual experiences of peer bullying victimization.

Furthermore, the field testing procedures used during the development of the contextual questionnaires help to ensure that the individual questions and resulting scales are reliable and valid measures of the desired constructs. After the field testing of the items, the Student Bullied at School Scale demonstrated a strong internal reliability, as measured by Cronbach's alpha, and each question contributed a portion of unique variance, which allows for insights into each student's overall experience of bullying victimization. Based on these measures of validity and reliability, the questions pertaining bullying victimization appear to demonstrate appropriate construct validity, and can be used for the further analyses.

Next, the construct validities of the different independent variables are considered. The independent variables used in this procedure are based on questions and scales that were developed and gathered in the same manner as the dependent variable scale. First, the four scales related to students level of confidence and enjoyment in math and science – Student Likes Learning Mathematics Scale, Student Confident in Mathematics Scale, Student Likes Learning Science Scale, and Student Confident in Science Scale – were all developed through a careful process of item development and field testing.

While the connection between confidence and enjoyment in mathematics and science and bullying victimization may not be obvious, there is literature to suggest that such a connection likely exists for female and minority students. In a study examining the rates of bullying victimization examined by academically successful children, Bishop and colleagues (2004) found that it was not necessarily GPA or academic success that was directly associated with bullying victimization, rather, it was the higher levels of academic effort and engagement that were directly related to higher rates of bullying victimization. The TIMSS 2011 does not directly ask students about the amount of time and effort he or she puts into mathematics and science study, but the assessment does ask students about their level of enjoyment and confidence in these two subjects.

Enjoyment and confidence are two factors that are strongly related to an individual developing intrinsic motivation (Ryan & Deci, 2000). Students who have demonstrated strong intrinsic motivation to complete a task are much more likely to be engaged and put forth more effort during that task than students who are not intrinsically motivated (Ryan & Deci, 2000). Therefore, by asking students to report their levels of enjoyment and confidence in mathematics, there is a strong reason to believe that students who enjoy and have more confidence in these two subjects are likely putting forth more effort in these subjects, and, therefore, are also at greater risk for experiencing bullying victimization.

Items for these scales assessing enjoyment and confidence display very strong face validity and asked students directly about their thoughts and feelings regarding the two subjects. For example, in the Student Likes Learning Mathematics [Science] Scale, students were asked to indicate how much they agreed with the statement, "I like mathematics [science]," which clearly assesses each students' enjoyment in the subject. Similarly, in the Student Confident in Mathematics [Science] Scale, students were asked to indicate how much they agreed with the statement, "I usually do well in the mathematics [science]," which accesses a student's level of self-confidence in the subject. Furthermore, these items were field tested and were shown to display appropriate internal reliability and factor loadings to each respective scale (Table 1). Also, each of these scales showed the desired positive relationship with measured student achievement on the mathematics and science assessments (Table 1). This evidence of strong concurrent validity with an objective measure provides further support for the overall construct validity of the scales.

Finally, the construct validity of the mathematics and science assessments should be considered. Items for both of these assessments were included after a lengthy, in-depth process that was discussed in detail above. A brief description of the development procedures vital to the construct validity of the measures will be discussed here. First, TIMSS 2011 developers pre-determined a set of frameworks and subject topics that were present in the vast majority of school curriculums and therefore considered to be the important topics to include in the assessments (Mullis & Martin, 2011). These frameworks were clearly related to those topics that are considered vital for mathematics and science success, and provided a firm foundation for appropriate construct validity.

In accordance with the established frameworks, explicit instructions were given to individuals familiar with writing mathematics and science assessment items (e.g., educators, educational consultants). These instructions were written in the *TIMSS 2011 Writing Guidelines* (Mullis & Martin, 2011) and ensured that each item matched the developed frameworks and desired question style. The TIMSS 2011 Development Teams and Committees then reviewed the items, rated the appropriateness of each question, and compiled the final assessments. Prior to the final assessments being compiled, however, all newly written questions were field tested with the help of over 60,000 students in 44 countries to determine the most appropriate items.

Finally, TIMSS 2011 developers were able to compile the final assessment packets from both newly written question items and question items that had been included in the previous iteration of the TIMSS assessment and had been vetted through the same process (Mullis & Martin, 2011). This whole process of item development and testing provides evidence that the TIMSS 2011 is a valid assessment of students' mathematics and science skills.

**Internal Validity**

Internal validity is a measure of how closely the differences between the variables measured in the research procedure match the actual differences that exist in the population, or occur during the real world situations (Gravetter & Forzano, 2009). There are several threats to internal validity, including confounding variables, assignment bias, history effects, maturation effects, and testing effects. Many of these threats to internal validity can be avoided through careful research design and sound theoretical basis.

One of the most difficult threats to internal validity to guard against is a confounding variable. Confounding variables are any effects that systematically impact the relationship between the dependent and independent variables and cause an alternative explanation for any observed relationships discovered between the variables (Gravetter & Forzano, 2009). In order to prevent confounding variables from impacting the current study, student and school factors that are not of interest to the research questions will be controlled for through the addition of those variables as covariates. Students' enjoyment, confidence, and skill in mathematics and science will be controlled for in the analysis so that a clearer picture of the relationship between gender, race, and bullying victimization can be examined.

Many of the other threats to internal validity can be avoided through carefully implemented research procedures and participant selection practices; each of which the TIMSS 2011 creators completed. The data gathered by the TIMSS 2011 procedure should not be impacted by history, maturation, or testing effects because the TIMSS 2011 data is collected only one time from each subject and the data collection period lasted only a few months, from April 2011 to June 2011, limiting the number of environmental factors that would be able to systematically affect a group of students. Similarly, as discussed previously in the discussion of construct validity, the instruments used in the TIMSS 2011 are based on strong theoretical foundations directly related to the constructs that the scales claim to measure, which further guards against the presence of other extraneous variables that would be responsible for the measured relationship between the dependent and independent variables.

One aspect of the TIMSS 2011 that does bring about some concern for the assessments' internal validity is the use of plausible values. As discussed previously, in order for the TIMSS 2011 to contain enough assessment items to satisfy the requirements of Item Response Theory (IRT) and allow the assessments enough statistical power to accurately measure student differences, the TIMSS 2011 contains such a large number of assessment items that expecting each student to complete each item would be unreasonable (Kastberg et al., 2013). Therefore, the TIMSS 2011 utilizes plausible values, which are estimates of how a student would have scored on the assessment items had they attempted each item based on a series of statistical calculations.

While this process of using plausible values is not unique to the TIMSS 2011 (see Beaton & González, 1995), its use does introduce more measurement error into the assessment procedure. Since each student is not answering every question, or even most of the questions, using the student's score as a measure of his or her achievement seems dubious. However, based on the proven statistical methods and the fact that researchers are instructed to conduct the analyses five separate times and take an average of these data to obtain the most accurate estimate of a student's true score, this use of plausible values should not impact the validity of the results.

**External Validity**

External validity is a measure of how applicable and generalizable research findings from a measured sample are to the population at large (Gravetter & Foranzo, 2009). Research is completed in order to draw conclusions about some characteristic or pattern of behavior in some target population. However, it is often impractical to study or survey an entire population of individuals. Therefore, researchers select a sample of individuals from the population in hopes that any sample characteristics discovered during the research matches, or is generalizable to, the characteristics of the entire population. But in order for researchers to claim that the research is applicable to the population, researchers must demonstrate that the sample studied matches the demographic, ideological, and other inherent qualities of the population. Studies whose samples accurately reflect their target populations, and therefore allow for accurate insights into that population, are considered to be externally valid (Gravetter & Foranzo, 2009).

The sampling method used by the TIMSS 2011 limits the threats to external validity. First, TIMSS 2011 developers were able to select the participating schools through random selection. Because developers had access to a list of all eligible public and private schools in the nation thanks to NCES databases, developers were able to randomly select from the entire population, which is considered the most valid method of selection (Gravetter & Foranzo, 2009). In order to further ensure that the measured sample matched the overall population, developers used census and economic data to create groups of schools that matched each criterion and selected schools in such a manner so that the sample proportions of each criteria corresponded with population proportions (Kastberg et al., 2013). This procedure is unlike most other large scale datasets that typically rely on complicated sampling procedures, such as oversampling minority populations or clustering students according to their presence in a larger institution or organization, to ensure that the gathered sample contained enough minority subjects to achieve a level of statistical power (Thomas & Heck, 2001). This sampling practice often leads to external validity issues because the sample characteristic proportions do not match the proportions of the population (Thomas & Heck, 2001). Due to this careful sampling procedure, the results of the TIMSS 2011 study should match the relationships and the patterns of behavior displayed by the general population and, therefore, display appropriate external validity.

**Procedure**

In order to obtain the necessary data, an application for the restricted use license for the TIMSS 2011 dataset was sent to the National Center of Education Statistics (NCES). The restricted use version of the dataset was chosen necessary in order to access the ethnic/racial demographic data of the sample. In accordance with NCES rules and regulations, a secure computer was purchased and stored in a private location. After ensuring the data's security, NCES approved the application for the TIMSS 2011 restricted use data set (License Control Number 15050027). Once the data had been received, it was stored securely and analyzed on a computer that was disconnected from any network communication at Duquesne University.

**Data Analysis**

Data from the TIMSS 2011 dataset was analyzed using IBM's SPSS Statistics software to determine the ability of the independent variables to predict changes in the dependent variable. While this research study consists of four separate research questions, each of the questions is structured the same way with the only difference between the questions being one independent variable (e.g., mathematics vs. science; gender status vs. racial minority status). Due to these structural similarities between research questions, the same analyses and statistical procedures will be completed for each question. Therefore, a detailed description of these procedures will be offered for the first research question and hypothesis, and then will be summarized for each of the subsequent questions.

The research questions and associated hypotheses are as follows:

Research Question 1: Does a student's identification with a specific race impact his or her experience of peer bullying victimization after controlling for the student's ability, enjoyment, and confidence in mathematics?

Hypothesis 1: After controlling for his or her mathematics ability, confidence, and enjoyment, a racial minority student will report higher frequencies of peer bullying victimization than a White student.

To answer this research question, an Analysis of Covariance (ANCOVA) will be completed. The ANCOVA procedure measures the effects of an independent variable on a dependent variable after equating the independent variable groups based on their performance or status of some covariate (Field, 2013). For this analysis, Peer Bullying Victimization will serve as the dependent variable, Race will serve as the independent or grouping variable, and Mathematics Ability, Student's Enjoyment in Mathematics, and Student's Confidence in Mathematics will serve as the covariates. Therefore, by using an ANCOVA, the differences between the amounts of bullying victimization experienced by students of each race will be determined after equalizing the mathematics performance, confidence, and enjoyment of all the students from each race.

Using an ANCOVA procedure also improves the researcher's ability to draw accurate conclusions from the gathered data. First, the ANCOVA reduces the amount of error variance between the independent variable groups and therefore increases the sensitivity of the $F$ statistic that determines if actual differences exist between the groups. This is accomplished through the addition of the covariates that account for the predictable error that occurs between the groups (Mertler & Vannatta, 2009).

Second, the ANCOVA assists in making accurate decisions in non-experimental studies, such as this one. This research question examines the inherent groups of racial status and not groups to which participants were randomly assigned. Since the study uses inherent groups, the groups likely differ on the covariates of mathematics performance, confidence, and enjoyment. By using the covariates, however, the effects of those covariates are partialed out, making the groups more equal (Mertler & Vannatta, 2009).

In order to appropriately complete an ANCOVA procedure, several assumptions must be met (Mertler & Vannatta, 2009). First, the responses gathered from each individual must be independent of each other. Second, scores on the dependent variable must be normally distributed. Third, the variances of the distributions of the scores on the dependent variable must be equal. Fourth, there must be a linear relationship between the dependent variable and the covariates. Fifth, there must be homogeneity between the regression slops of the covariates. Each of these five assumptions will be tested and results are listed and discussed below. Finally, the last assumption is that the covariate has been reliably measured. The TIMSS 2011 should be considered to meet this assumption based on the theoretically based and systematic manner in which the TIMSS 2011 mathematics assessment was developed and administered to students.

For this research question, the null and alternative hypotheses and alpha level necessary to reject the null hypothesis are as follows:

$H_0$ = there is no difference between the bullying rates of racial minority students and White students ($\mu_0 = \mu_1$)

$H_1$ = racial minority students experience significantly higher rates of bullying victimization than White students ($\mu_0 < \mu_1$; $\alpha < .05$)

Research Question 2: Does a student's identification with a specific race impact his or her experience of peer bullying victimization after controlling for the student's ability, enjoyment, and confidence in science?

Hypothesis 1: After controlling for his or her science ability, confidence, and enjoyment, a racial minority student will report higher frequencies of peer bullying victimization than a White student.

To answer this research question, another ANCOVA procedure will be completed. The reasoning behind using this analysis, as well as the consideration of the assumptions associated with the analysis, matches the discussion above. For this analysis, Peer Bullying Victimization will serve as the dependent variable. Race will serve as the independent or grouping variable, and Science Ability, Student's Enjoyment in Science, and Student's Confidence in Science will serve as the covariates. Therefore, by using an ANCOVA, the differences between the amounts of bullying victimization experienced by students of each race will be determined after equalizing the science performance, enjoyment, and confidence of all the students from each race.

For this research question, the null and alternative hypotheses and alpha level necessary to reject the null hypothesis are as follows:

$H_0$ = there is no difference between the bullying rates of racial minority students and White students ($\mu_0 = \mu_1$)

$H_1$ = racial minority students experience significantly higher rates of bullying victimization than White students ($\mu_0 < \mu_1$; $\alpha < .05$)

Research Question 3: Does a student's gender impact his or her experience of peer bullying victimization after controlling for the student's ability, enjoyment, and confidence in mathematics?

Hypothesis 1: After controlling for her mathematics ability, confidence, and enjoyment, a female student will report higher frequencies of peer bullying victimization than a male student.

To answer this research question, another ANCOVA procedure will be completed. The reasoning behind using this analysis, as well as the consideration of the assumptions associated with the analysis, matches the discussion above. For this analysis, Peer Bullying Victimization will serve as the dependent variable, Gender will serve as the independent or grouping variable, and Mathematics Ability, Student's Enjoyment in Mathematics, and Student's Confidence in Mathematics will serve as the covariates. Therefore, by using an ANCOVA, the differences between the amounts of bullying victimization experienced by students of each gender will be determined after equalizing the mathematics performance, enjoyment, and confidence of all the students from each gender.

For this research question, the null and alternative hypotheses and alpha level necessary to reject the null hypothesis are as follows:

$H_0$ = there is no difference between the bullying rates of female students and male students ($\mu_0 = \mu_1$)

$H_1$ = female students experience significantly higher rates of bullying victimization than male students ($\mu_0 < \mu_1$; $\alpha < .05$)

Research Question 4: Does a student's gender impact his or her experience of peer bullying victimization after controlling for the student's ability, enjoyment, and confidence in science?

Hypothesis 1: After controlling for her science ability, confidence, and enjoyment, a female student will report higher frequencies of peer bullying victimization than a male student.

To answer this research question, another ANCOVA procedure will be completed. The reasoning behind using this analysis as well as the consideration of the assumptions associated with the analysis matches the discussion above. For this analysis, Peer Bullying Victimization will serve as the dependent variable, Gender will serve as the independent or grouping variable, and Science Ability, Student's Enjoyment in Science, and Student's Confidence in Science will serve as the covariates. Therefore, by using an ANCOVA, the differences between the amounts of bullying victimization experienced by students of each race/ethnicity will be determined after equalizing the science performance, enjoyment, and confidence of all the students from both genders.

For this research question, the null and alternative hypotheses and alpha level necessary to reject the null hypothesis are as follows:

$H_0$ = there is no difference between the bullying rates of female students and male students ($\mu_0 = \mu_1$)

$H_1$ = female students experience significantly higher rates of bullying victimization than male students ($\mu_0 < \mu_1$; $\alpha < .05$)

## Summary

In order to answer the four research questions posed in this study, information from the United States sample of the Trends in International Mathematics and Science Study of 2011 (TIMSS 2011) will be analyzed. The TIMSS 2011 was administered to 10,480 eighth grade students from a nationally representative sample of students across the United States. It consists of a series of questionnaires administered to students, teachers, and school administrators as well as standardized mathematics and science assessments. The TIMSS 2011 was developed in a careful, systematic, and theory-driven manner that provides the survey with strong construct, internal, and external validity. The data gathered from the TIMSS 2011 restricted use dataset will be analyzed using an Analysis of Covariance (ANCOVA) statistical procedure to determine any relationships between an individual's self-report of peer bullying victimization and that individual's gender and racial minority status after controlling for his or her mathematics and science ability, enjoyment, and confidence.

# CHAPTER 4 – Results

The following information describes the process in which the data from the TIMSS 2011 restricted use dataset was prepared for analysis. This chapter will also illustrate how the assumptions that accompany the completion of an Analysis of Covariance (ANCOVA) were checked. Finally, the results of several preliminary analyses, as well as the main ANCOVA analyses will be presented.

## Data Preparation

The first step in the analysis process was to ensure that the data used was complete and allowed for the most accurate analyses. The original sample size consisted of 10,480 participants and the demographic frequencies of each of the independent, dependent, and covariate variables, as well as the original number of missing values are presented in Table 4. Based on the NCES regulations, all unweighted sample sizes will be rounded to the nearest 10 to ensure participant anonymity.

As shown in Table 4, the TIMSS 2011 dataset was remarkably complete, as the variable with the most missing data, the Student Confidence with Science Scale, was only missing 1.9% of data points. In order to account for the small amount of missing data, the first step was to delete the 60 individuals who had not been administered the Student Questionnaire, and, therefore, were missing data for each of the Student Questionnaire variable scales used. The next step was to delete an additional 20 students who were administered the Student Questionnaire, but whose response pattern made it impossible to calculate any of the variable scales that were of interest to this study. The next step was to delete an additional 40 individual students who had not reported their racial demographic information. These deletions resulted in a total of 130 cases being removed

Table 4

*Sample Size and Number of Missing Cases in Original TIMSS 2011 Data Set for Each Variable Used*

| Variable | | Sample Size | % of Data |
|---|---|---|---|
| Gender | | | |
| | Female | 5,300 | 51 |
| | Male | 5,180 | 49 |
| | Omitted | 0 | 0 |
| | Not Administered | 0 | 0 |
| | Total Missing | 0 | 0 |
| Race | | | |
| | White | 5,180 | 49 |
| | Black | 1,260 | 12 |
| | Hispanic | 2,690 | 26 |
| | Asian | 480 | 5 |
| | Native American | 120 | 1 |
| | Pacific Islander | 100 | 1 |
| | 2 or more races | 550 | 5 |
| | Omitted or Invalid | 110 | 1 |
| Student Bullied at School Scale | | | |
| | Complete | 10,390 | 9817 |
| | Omitted | 80 | 0.7 |
| | Not Administered | 60 | 0.6 |
| | Total Missing | 140 | 1.3 |
| Student Bullied at School Index | | | |
| | Complete | 10,390 | 98.7 |
| | Omitted | 80 | 0.7 |
| | Not Administered | 60 | 0.6 |
| | Total Missing | 140 | 1.3 |
| Student Likes Learning Math Scale | | | |
| | Complete | 10,340 | 98.7 |
| | Omitted | 80 | 0.7 |
| | Not Administered | 60 | 0.6 |
| | Total Missing | 140 | 1.3 |
| Student Confident in Math Scale | | | |
| | Complete | 10,330 | 98.6 |
| | Omitted | 80 | 0.8 |
| | Not Administered | 60 | 0.6 |

| | | |
|---|---|---|
| Total Missing | 140 | 1.4 |
| Student Likes Learning Science Scale | | |
| Complete | 10,310 | 98.4 |
| Omitted | 110 | 1.0 |
| Not Administered | 60 | 0.6 |
| Total Missing | 170 | 1.6 |
| Student Confident in Science Scale | | |
| Complete | 10,280 | 98.1 |
| Omitted | 130 | 1.3 |
| Not Administered | 60 | 0.6 |
| Total Missing | 200 | 1.9 |
| Math Assessment Plausible Values 1-5 | | |
| Complete | 10,480 | 100 |
| Omitted | 0 | 0 |
| Not Administered | 0 | 0 |
| Total Missing | 0 | 0 |
| Science Assessment Plausible Values 1-5 | | |
| Complete | 10,480 | 100 |
| Omitted | 0 | 0 |
| Not Administered | 0 | 0 |
| Total Missing | 0 | 0 |

Note: $N$ = 10,480

listwise from the original dataset, which accounted for 1.3% of the original dataset. After the cases were removed, the final sample size equaled 10,350 participants. Table 5 illustrates the number of completed and missing cases for each variable used following the case deletions.

After the deletions, each of the scales intended for use contained 1.1% or fewer missing data points. Roth (1994) suggests that when using any dataset that contains less than 10% missing data, it is appropriate to replace missing data points with the series mean. When used with a dataset in which missing data accounts for 10% or more of the available data points, this basic method of missing data imputation can lead data to

Table 5

*Sample Size and Missing Cases for each Variable Used Following Case Deletions*

| Variable | | Sample Size | % of Data |
|---|---|---|---|
| Gender | | | |
| | Female | 5,240 | 51 |
| | Male | 5,110 | 49 |
| | Total Missing | 0 | 0 |
| Race | | | |
| | White | 5,170 | 50 |
| | Black | 1,250 | 12 |
| | Hispanic | 2,680 | 26 |
| | Asian | 480 | 5 |
| | Native American | 120 | 1 |
| | Pacific Islander | 100 | 1 |
| | 2 or more races | 550 | 5 |
| | Total Missing | 0 | 0 |
| Student Bullied at School Scale | | | |
| | Complete | 10,300 | 99.5 |
| | Total Missing | 50 | 0.5 |
| Student Bullied at School Index | | | |
| | Complete | 10,300 | 99.5 |
| | Total Missing | 50 | 0.5 |
| Student Likes Learning Math Scale | | | |
| | Complete | 10,300 | 99.5 |
| | Total Missing | 50 | 0.5 |
| Student Confident in Math Scale | | | |
| | Complete | 10,270 | 99.2 |
| | Total Missing | 80 | 0.8 |
| Student Likes Learning Science Scale | | | |
| | Complete | 10,290 | 99.5 |
| | Total Missing | 60 | 0.5 |
| Student Confident in Science Scale | | | |
| | Complete | 10,240 | 98.9 |
| | Total Missing | 110 | 1.1 |
| Math Assessment Plausible Values 1-5 | | | |

| | | |
|---|---|---|
| Complete | 10,350 | 100 |
| Total Missing | 0 | 0 |
| Science Assessment Plausible Values 1-5 | | |
| Complete | 10,350 | 100 |
| Total Missing | 0 | 0 |

Note: $N$ = 10,350

become more peaked and cause the sample to fail to be appropriately representative of the measured population. However, when used with small proportions of missing data, using the series mean as the missing data value will make little statistical impact on the data while maximizing the available sample size (Roth, 1994). Due to the small percentage of missing data points, the series mean for each of the scales was calculated and imputed for the missing data points. As a result, the final sample size used in all subsequent analyses is 10,350 students.

**Demographic Data**

After establishing the final sample, descriptive statistics were calculated for each demographic group on each of the scales used as variables during the subsequent analyses. Table 6 presents the descriptive statistics for the Student Questionnaire sales, Table 7 presents the descriptive statistics for the Mathematics Plausible Values, and Table 8 presents the descriptive statistics for the Science Plausible Values.

Table 6

*Mean Responses to Student Questionnaire Question Items by Gender and Race*

*Demographics of Student Sample*

| Demographic (N) | Student Bullied at School Scale | | Student Likes Learning Math Scale | | Student Likes Learning Science Scale | | Student Confident Learning Math Scale | | Student Confident Learning Science Scale | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Gender | | | | | | | | | |
| | *M* | *SD* | *M* | *SD* | *M* | *SD* | *M* | *SD* | *M* | *SD* |
| Female (5,240) | 10.15 | 1.83 | 9.43 | 2.01 | 9.40 | 2.04 | 10.36 | 2.29 | 10.13 | 2.13 |
| Male (5,110) | 10.07 | 2.01 | 9.58 | 2.06 | 9.77 | 2.07 | 10.76 | 2.22 | 10.56 | 2.11 |
| | Race/Ethnicity | | | | | | | | | |
| | *M* | *SD* | *M* | *SD* | *M* | *SD* | *M* | *SD* | *M* | *SD* |
| White (5,170) | 10.04 | 1.90 | 9.32 | 2.05 | 9.64 | 2.08 | 10.62 | 2.29 | 10.54 | 2.14 |
| Black (1,250) | 10.10 | 1.94 | 9.98 | 1.98 | 9.39 | 2.07 | 10.65 | 2.26 | 10.11 | 2.23 |
| Hisp. (2,680) | 10.29 | 1.93 | 9.52 | 1.99 | 9.45 | 2.00 | 10.27 | 2.19 | 9.98 | 1.96 |
| Asian (480) | 10.22 | 1.89 | 10.16 | 2.03 | 10.06 | 1.99 | 11.15 | 2.24 | 10.60 | 2.15 |
| Native Am. (120) | 9.85 | 2.33 | 9.44 | 2.22 | 9.59 | 1.91 | 10.34 | 2.38 | 10.11 | 2.09 |
| Pacific Island. (100) | 10.18 | 1.92 | 9.42 | 2.17 | 9.50 | 1.94 | 10.19 | 2.10 | 10.14 | 2.11 |
| Two or More Races (550) | 9.89 | 1.98 | 9.57 | 1.96 | 9.68 | 2.20 | 10.70 | 2.22 | 10.55 | 2.27 |
| Total | 10.11 | 1.92 | 9.50 | 2.04 | 9.58 | 2.06 | 10.56 | 2.26 | 10.34 | 2.13 |

Note: $N$ = 10,350

Table 7

*Mean Scores on each of the Mathematics Plausible Values by Gender and Race*

*Demographics of Student Sample*

| Demo-graphic (N) | Plausible Value #1 | | Plausible Value #2 | | Plausible Value #3 | | Plausible Value #4 | | Plausible Value #5 | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Gender | | | | | | | | | |
| | *M* | *SD* | *M* | *SD* | *M* | *SD* | *M* | *SD* | *M* | *SD* |
| Female (5,240) | 508.6 | 74.1 | 508.9 | 75.2 | 510.1 | 75.2 | 509.8 | 75.1 | 508.2 | 75.1 |
| Male (5,110) | 511.2 | 76.8 | 512.2 | 77.2 | 512.5 | 77.3 | 512.4 | 76.7 | 512.9 | 77.0 |
| | Race/Ethnicity | | | | | | | | | |
| | *M* | *SD* | *M* | *SD* | *M* | *SD* | *M* | *SD* | *M* | *SD* |
| White (5,170) | 529.4 | 69.4 | 530.1 | 69.9 | 531.0 | 70.0 | 530.7 | 70.1 | 529.8 | 70.2 |
| Black (1,250) | 469.2 | 72.5 | 469.8 | 73.2 | 469.3 | 74.9 | 470.2 | 73.6 | 470.5 | 73.3 |
| Hisp. (2,680) | 483.7 | 69.8 | 484.3 | 70.6 | 485.0 | 70.0 | 485.0 | 69.3 | 484.3 | 69.9 |
| Asian (480) | 562.6 | 80.6 | 565.4 | 79.8 | 564.9 | 79.6 | 565.6 | 81.5 | 565.8 | 82.6 |
| Native Am. (120) | 465.8 | 72.8 | 460.9 | 72.5 | 464.1 | 73.2 | 467.8 | 72.8 | 465.2 | 70.9 |
| Pacific Island. (100) | 482.4 | 63.3 | 481.9 | 71.6 | 487.6 | 67.7 | 490.9 | 62.0 | 485.2 | 65.1 |
| Two or More Races (550) | 515.7 | 70.6 | 515.0 | 72.3 | 517.4 | 71.7 | 513.2 | 71.2 | 515.0 | 71.8 |
| Total | 509.9 | 75.4 | 510.5 | 76.2 | 511.3 | 76.2 | 511.1 | 75.9 | 510.5 | 76.1 |

Note: $N = 10{,}350$

Table 8

*Mean Scores on each of the Science Plausible Values by Gender and Race Demographics of Student Sample*

| Demographic (N) | Plausible Value #1 | | Plausible Value #2 | | Plausible Value #3 | | Plausible Value #4 | | Plausible Value #5 | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Gender | | | | | | | | | |
| | *M* | *SD* | *M* | *SD* | *M* | *SD* | *M* | *SD* | *M* | *SD* |
| Female (5,240) | 519.9 | 78.0 | 518.8 | 77.7 | 519.4 | 77.2 | 519.3 | 78.2 | 520.5 | 77.5 |
| Male (5,110) | 530.7 | 83.4 | 530.6 | 81.9 | 531.2 | 81.4 | 530.7 | 81.7 | 531.8 | 83.0 |
| | Race/Ethnicity | | | | | | | | | |
| | *M* | *SD* | *M* | *SD* | *M* | *SD* | *M* | *SD* | *M* | *SD* |
| White (5,170) | 552.7 | 71.3 | 551.7 | 70.7 | 552.2 | 70.2 | 551.8 | 71.1 | 552.7 | 71.0 |
| Black (1,250) | 475.3 | 78.1 | 474.4 | 78.3 | 474.7 | 77.2 | 475.2 | 76.0 | 475.7 | 78.0 |
| Hisp. (2,680) | 492.4 | 76.9 | 493.1 | 74.3 | 493.7 | 74.4 | 492.7 | 75.2 | 493.5 | 75.4 |
| Asian (480) | 552.8 | 80.0 | 549.2 | 82.6 | 552.1 | 80.1 | 553.4 | 82.5 | 551.6 | 82.8 |
| Native Am. (120) | 486.3 | 78.2 | 483.2 | 74.9 | 484.6 | 74.9 | 481.6 | 81.5 | 484.7 | 78.3 |
| Pacific Island. (100) | 476.3 | 71.2 | 480.3 | 75.9 | 478.5 | 73.1 | 479.3 | 73.1 | 476.0 | 74.5 |
| Two or More Races (550) | 534.3 | 73.1 | 534.1 | 74.5 | 534.5 | 73.6 | 535.3 | 73.6 | 536.5 | 71.7 |
| Total | 525.2 | 80.9 | 524.6 | 80.0 | 525.2 | 79.5 | 524.9 | 80.2 | 526.1 | 80.5 |

Note: *N* = 10,350

## Assumption Testing

In order to ensure the accurate interpretation of completed analyses, data must meet a set of assumptions. These assumptions ensure that the statistical differences found in the data are not unduly caused by the data's distribution or other qualities that over- or under-inflate results. However, due to the robust nature of the ANCOVA procedure and the large sample size gathered by the TIMSS 2011, many of the assumptions that are associated with significance testing using linear models, in general, and the use of ANCOVA methodology, specifically, are shown to have less of an impact on the outcomes of the analyses (Porter & Raudenbush, 1987). Despite this consideration, the assumptions made for an ANCOVA as described by Field (2013) are listed and discussed below.

### Independence of Covariates and Independent Variables

The first assumption when completing an ANCOVA is related to whether or not the covariates and the treatment variables are independent of each other. If the covariates and the treatment variables are not independent, the two variables will likely account for a shared portion of the explained variance illustrated by the ANCOVA. This would impact the ability to accurately determine the impact of the independent variable on the dependent variable (Field, 2013). In order to determine if the treatment variables of gender and racial demographic show significant differences related to the responses or abilities measured by the TIMSS 2011 scales, a series of analysis of variance (ANOVA) analyses were completed using the Gender or Race variable as the independent variable, and each of the originally intended covariate scales as the dependent variable.

Table 9

*ANOVA Results for Comparisons between Demographic Categories and Responses to Student Questionnaire Items*

| Scale | | Sum of Squares | Df | Mean Square | F | Sig |
|---|---|---|---|---|---|---|
| | | Gender | | | | |
| Student Bullied at School Scale | Between Groups | 18.16 | 1 | 18.16 | 4.91 | .027 |
| | Within Groups | 38270.69 | 10350 | 3.70 | | |
| | Total | 38288.85 | 10350 | | | |
| Student Likes Learning Math Scale | Between Groups | 63.61 | 1 | 63.61 | 15.35 | .000 |
| | Within Groups | 42885.14 | 10350 | 4.15 | | |
| | Total | 42948.75 | 10350 | | | |
| Student Likes Learning Science Scale | Between Groups | 358.09 | 1 | 358.09 | 85.03 | .000 |
| | Within Groups | 43571.91 | 10350 | 4.21 | | |
| | Total | 43930.00 | 10350 | | | |
| Student Confident in Math Scale | Between Groups | 400.65 | 1 | 400.65 | 78.84 | .000 |
| | Within Groups | 52574.44 | 10350 | 5.08 | | |
| | Total | 52975.09 | 10350 | | | |
| Student Confident in Science Scale | Between Groups | 474.58 | 1 | 474.58 | 105.63 | .000 |
| | Within Groups | 46481.63 | 10350 | 4.49 | | |
| | Total | 46956.20 | 10350 | | | |
| | | Race/Ethnicity | | | | |
| Student Bullied at School Scale | Between Groups | 151.33 | 6 | 25.22 | 6.84 | .000 |
| | Within Groups | 38137.52 | 10340 | 3.69 | | |
| | Total | 38288.85 | 10350 | | | |
| Student Likes Learning Math Scale | Between Groups | 669.08 | 6 | 111.51 | 27.27 | .000 |
| | Within Groups | 42279.67 | 10340 | 4.09 | | |
| | Total | 42948.75 | 10350 | | | |
| Student Likes Learning Science Scale | Between Groups | 223.46 | 6 | 37.24 | 8.812 | .000 |
| | Within Groups | 43706.54 | 10340 | 4.23 | | |
| | Total | 43930.00 | 10350 | | | |
| Student Confident in Math Scale | Between Groups | 445.14 | 6 | 74.19 | 14.61 | .000 |
| | Within Groups | 52529.95 | 10340 | 5.08 | | |
| | Total | 52975.09 | 10350 | | | |
| Student Confident in Science Scale | Between Groups | 694.43 | 6 | 115.74 | 25.87 | .000 |
| | Within Groups | 46261.78 | 10340 | 4.74 | | |
| | Total | 46956.20 | 10350 | | | |

Table 10

*Mean Difference Results of Bonferroni-Adjusted Pairwise Post-Hoc Analyses between Students' Race and Responses on the Student Questionnaire Scales*

| Scale | (I) Race | (J) Race | Mean Difference (I-J) | Std. Error | Sig. |
|---|---|---|---|---|---|
| Student Bullied at School Scale | White | Black | -.062 | .060 | 1.00 |
| | | Hispanic | -.249 | .046 | .000 |
| | | Asian | -.184 | .092 | .960 |
| | | Native American | .186 | .179 | 1.00 |
| | | Pacific Islander | -.138 | .196 | 1.00 |
| | | 2 or more races | .144 | .086 | 1.00 |
| | Black | Hispanic | -.187 | .066 | .092 |
| | | Asian | -.122 | .103 | 1.00 |
| | | Native American | .248 | .185 | 1.00 |
| | | Pacific Islander | -.076 | .201 | 1.00 |
| | | 2 or more races | .206 | .098 | .752 |
| | Hispanic | Asian | .065 | .096 | 1.00 |
| | | Native American | .436 | .181 | .334 |
| | | Pacific Islander | .111 | .198 | 1.00 |
| | | 2 or more races | .394 | .090 | .000 |
| | Asian | Native American | .370 | .198 | 1.00 |
| | | Pacific Islander | .045 | .213 | 1.00 |
| | | 2 or more races | .328 | .120 | .134 |
| | Native American | Pacific Islander | -.324 | .262 | 1.00 |
| | | 2 or more races | -.042 | .195 | 1.00 |
| | Pacific Islander | 2 or more races | .282 | .211 | 1.00 |
| Student Likes Learning Math Scale | White | Black | -.660 | .064 | .000 |
| | | Hispanic | -.206 | .048 | .000 |
| | | Asian | -.841 | .097 | .000 |
| | | Native American | -.119 | .188 | 1.00 |
| | | Pacific Islander | -.107 | .206 | 1.00 |
| | | 2 or more races | -.256 | .103 | .099 |
| | Black | Hispanic | .454 | .069 | .000 |
| | | Asian | -.182 | .109 | 1.00 |
| | | Native American | .541 | .195 | .115 |
| | | Pacific Islander | .553 | .212 | .191 |

| | | | | | |
|---|---|---|---|---|---|
| | | 2 or more races | .404 | .103 | .002 |
| | Hispanic | Asian | -.635 | .101 | .000 |
| | | Native American | .087 | .190 | 1.00 |
| | | Pacific Islander | .100 | .208 | 1.00 |
| | | 2 or more races | -.050 | .095 | 1.00 |
| | Asian | Native American | .723 | .208 | .011 |
| | | Pacific Islander | .735 | .224 | .022 |
| | | 2 or more races | .585 | .127 | .000 |
| | Native American | Pacific Islander | .013 | .276 | 1.00 |
| | | 2 or more races | -.137 | .205 | 1.00 |
| | Pacific Islander | 2 or more races | -.150 | .222 | 1.00 |
| Student Likes Learning Science Scale | White | Black | .251 | .065 | .002 |
| | | Hispanic | .188 | .049 | .003 |
| | | Asian | -.419 | .099 | .000 |
| | | Native American | .053 | .191 | 1.00 |
| | | Pacific Islander | .149 | .210 | 1.00 |
| | | 2 or more races | -.038 | .092 | 1.00 |
| | Black | Hispanic | -.063 | .070 | 1.00 |
| | | Asian | -.670 | .111 | .000 |
| | | Native American | -.199 | .198 | 1.00 |
| | | Pacific Islander | -.102 | .216 | 1.00 |
| | | 2 or more races | -.289 | .105 | .125 |
| | Hispanic | Asian | -.607 | .102 | .000 |
| | | Native American | -.136 | .193 | 1.00 |
| | | Pacific Islander | -.039 | .211 | 1.00 |
| | | 2 or more races | -.226 | .096 | .396 |
| | Asian | Native American | .471 | .211 | .544 |
| | | Pacific Islander | .568 | .228 | .269 |
| | | 2 or more races | .381 | .129 | .065 |
| | Native American | Pacific Islander | .097 | .281 | 1.00 |
| | | 2 or more races | -.090 | .209 | 1.00 |
| | Pacific Islander | 2 or more races | -.187 | .225 | 1.00 |
| Student Confident Learning Math Scale | White | Black | -.018 | .071 | 1.00 |
| | | Hispanic | .352 | .054 | .000 |
| | | Asian | -.523 | .108 | .000 |
| | | Native American | .282 | .210 | 1.00 |
| | | Pacific Islander | .435 | .230 | 1.00 |

| | | | | | |
|---|---|---|---|---|---|
| | | 2 or more races | -.071 | .101 | 1.00 |
| | Black | Hispanic | .370 | .077 | .000 |
| | | Asian | -.505 | .121 | .001 |
| | | Native American | .309 | .217 | 1.00 |
| | | Pacific Islander | .454 | .236 | 1.00 |
| | | 2 or more races | -.053 | .115 | 1.00 |
| | Hispanic | Asian | -.875 | .112 | .000 |
| | | Native American | -.070 | .212 | 1.00 |
| | | Pacific Islander | .083 | .232 | 1.00 |
| | | 2 or more races | -.423 | .105 | .001 |
| | Asian | Native American | .806 | .232 | .011 |
| | | Pacific Islander | .959 | .250 | .003 |
| | | 2 or more races | .452 | .141 | .029 |
| | Native American | Pacific Islander | .153 | .308 | 1.00 |
| | | 2 or more races | -.353 | .229 | 1.00 |
| | Pacific Islander | 2 or more races | -.507 | .247 | .846 |
| Student Confident in Science Scale | White | Black | .431 | .067 | .000 |
| | | Hispanic | .564 | .050 | .000 |
| | | Asian | -.060 | .101 | 1.00 |
| | | Native American | .423 | .197 | .664 |
| | | Pacific Islander | .398 | .216 | 1.00 |
| | | 2 or more races | -.005 | .095 | 1.00 |
| | Black | Hispanic | .133 | .072 | 1.00 |
| | | Asian | -.491 | .114 | .000 |
| | | Native American | .008 | .204 | 1.00 |
| | | Pacific Islander | -.033 | .222 | 1.00 |
| | | 2 or more races | -.436 | .108 | .001 |
| | Hispanic | Asian | -.624 | .105 | .000 |
| | | Native American | -.141 | .199 | 1.00 |
| | | Pacific Islander | -.166 | .218 | 1.00 |
| | | 2 or more races | -.570 | .099 | .000 |
| | Asian | Native American | .483 | .218 | .553 |
| | | Pacific Islander | .458 | .245 | 1.00 |
| | | 2 or more races | .054 | .132 | 1.00 |
| | Native American | Pacific Islander | -.025 | .289 | 1.00 |
| | | 2 or more races | -.429 | .214 | .959 |
| | Pacific Islander | 2 or more races | -.403 | .231 | 1.00 |

**Differences in responses to Student Questionnaire scales.** The first set of ANOVA analyses measured the gender and racial differences measured by the scales on the Student Questionnaire. Those results are presented in Table 9. If a significant difference was found by the ANOVA analysis when using the Race variable as the independent variable, then Bonferroni-adjusted post-hoc analyses were completed in order to determine which pairwise group differences were most responsible for the overall significant difference between racial demographic groups. Those results are presented in Table 10.

Results of these analyses indicate that responses on the Student Questionnaire frequently and significantly differ according to students' gender and racial demographics. On the Student Bullied at School Scale, female and male students' responses were significantly different, $F(1, 10350) = 4.91$; $p = .03$, with female students reporting more frequent experiences of bullying. Responses on this scale also differed by students' race, $F(6, 10340) = 25.22$; $p < .01$, with Hispanic students reporting significantly higher rates of bullying experiences than White and Multi-racial students.

On the Student Likes Learning Math Scale, female and male students were significantly different, $F(1, 10350) = 15.35$; $p < .01$, with male students reporting higher rates of math enjoyment. Responses also differed significantly by race, $F(6, 10340) = 27.27$; $p = .03$. Asian students reported the highest rates of math enjoyment, and those ratings were significantly higher than the reports made by White, Hispanic, Native American, Pacific Islander, and Multi-racial students. Black students reported the next highest ratings of enjoyment with math, and those students' ratings were significantly higher than White, Hispanic, and Multi-racial students' reports. Finally, Hispanic students reported the next highest level of enjoyment with math, and their reports were significantly higher than White students' reports of math enjoyment.

On the Student Likes Learning Science scale, responses, again, differed by gender, $F(1, 10350) = 85.03$; $p < .01$, with male students reporting significantly higher rates of enjoyment with science. Responses also differed by race, $F(6, 10340) = 8.81$; $p < .01$. Asian students reported significantly higher levels of science enjoyment than did White, Black, and Hispanic students, while White students reported significantly higher levels of enjoyment than did Black and Hispanic students.

On the Student Confident in Math scale, responses continued to differ by gender and by race. The gender differences were statistically significant, $F(1, 10350) = 78.84$; $p < .01$, with male students reporting significantly more confidence in math than their female peers. The racial differences were also statistically significant, $F(6, 10340) = 14.61$; $p < .01$. Similar to the results found related to math enjoyment, Asian students reported significantly higher rates of math confidence than White, Black, Hispanic, Native American, Pacific Islander, and Multi-racial students. Results also indicate that both White and Black students reported higher rates of math confidence than did Hispanic students.

Finally, on the Student Confident in Science scale, significant gender differences continued, $F(1, 10350) = 105.63$; $p < .01$, as male students reported higher rates of confidence with science than did female students. Significant racial differences persisted as well, $F(6, 10340) = 25.87$; $p < .01$, as Asian, White, and Multi-racial students each reported higher rates of confidence with science than both Black and Hispanic students.

These findings indicate that results obtained by each of the Student Questionnaire scales that are intended to serve as covariates during the ANCOVA analysis do vary significantly across both treatment groups of gender and racial demographics. As discussed by Field (2013), this lack of independence between the covariates and the treatment groups will complicate the interpretation of the ANCOVA findings as both the covariates and the independent variable will account for a shared portion of the explained variance in the Peer Bullying Victimization outcome variable.

**Differences in scores on the Math Assessment Plausible Values.** In order to determine if gender and/or racial demographic differences also exist in terms of mathematics abilities as measured by the TIMSS 2011 Math Assessments, ANOVA analyses were completed to gauge the differences between each of the developed Plausible Value scores. Because each Plausible Value is calculated using different possible score patterns, it is important to consider each Plausible Value separately. However, when interpreting the differences between the math Plausible Values earned by one demographic group compared to another, the more values that are shown to be significantly different, the stronger the conclusion that those two demographic groups do, in reality, differ in mathematics ability. The ANOVA results between demographic groups in terms of math abilities are presented in Table 11, while Table 12 presents the Bonferroni-adjusted pairwise comparisons between racial demographic groups.

Performance on the TIMSS 2011 Math Assessment appears to vary significantly according to students' racial demographics; however, variability related to gender is somewhat less stable. When the differences in math performance between male and female students are examined, male students are shown to earn significantly higher scores

Table 11

*ANOVA Results for Comparisons between Demographic Categories and Mathematics Assessment*

| Scale | | Sum of Squares | Df | Mean Square | F | Sig |
|---|---|---|---|---|---|---|
| | | Gender | | | | |
| Plausible Value #1 | Between Groups | 17836.4 | 1 | 17836.4 | 3.13 | .077 |
| | Within Groups | 58873871.8 | 10350 | 5690.5 | | |
| | Total | 58891708.1 | 10350 | | | |
| Plausible Value #2 | Between Groups | 28457.7 | 1 | 28457.7 | 4.90 | .027 |
| | Within Groups | 60061278.8 | 10350 | 5805.3 | | |
| | Total | 60089736.6 | 10350 | | | |
| Plausible Value #3* | Between Groups | 14386.8 | 1 | 14386.8 | 2.48 | .116 |
| | Within Groups | 60128582.8 | 10350 | 5811.8 | | |
| | Total | 60142969.6 | 10350 | | | |
| Plausible Value #4 | Between Groups | 17474.7 | 1 | 17474.7 | 3.04 | .081 |
| | Within Groups | 59547474.3 | 10350 | 5755.6 | | |
| | Total | 59564949.0 | 10350 | | | |
| Plausible Value #5* | Between Groups | 57851.4 | 1 | 57851.4 | 10.00 | .002 |
| | Within Groups | 59868030.3 | 10350 | 5786.6 | | |
| | Total | 59925881.6 | 10350 | | | |
| | | Race | | | | |
| Plausible Value #1* | Between Groups | 7526058.3 | 6 | 1245343.1 | 252.5 | .000 |
| | Within Groups | 51365649.8 | 10340 | 4967.2 | | |
| | Total | 58891708.1 | 10350 | | | |
| Plausible Value #2 | Between Groups | 7704016.9 | 6 | 1284002.8 | 253.5 | .000 |
| | Within Groups | 52385719.7 | 10340 | 5065.8 | | |
| | Total | 60089736.6 | 10350 | | | |
| Plausible Value #3 | Between Groups | 7771931.0 | 6 | 1295321.8 | 255.7 | .000 |
| | Within Groups | 52371038.6 | 10340 | 5064.4 | | |
| | Total | 60142969.6 | 10350 | | | |
| Plausible Value #4 | Between Groups | 7590617.0 | 6 | 1265102.8 | 251.7 | .000 |
| | Within Groups | 51974332.1 | 10340 | 5026.0 | | |
| | Total | 59564949.0 | 10350 | | | |
| Plausible Value #5 | Between Groups | 7544077.7 | 6 | 1257346.3 | 248.2 | .000 |
| | Within Groups | 52381804.0 | 10340 | 5065.4 | | |
| | Total | 59925881.6 | 10350 | | | |

Note: * denotes Plausible Value that will be used during subsequent analyses

Table 12

*Mean Difference Results of Bonferroni-Adjusted Pairwise Post-Hoc Analyses between Students' Race and TIMSS 2011 Math Assessment Plausible Values*

| Plausible Value | (I) Race | (J) Race | Mean Difference (I-J) | Std. Error | Sig. |
|---|---|---|---|---|---|
| Math Plausible Value #1 | White | Black | 60.21 | 2.22 | .000 |
| | | Hispanic | 45.69 | 1.68 | .000 |
| | | Asian | -33.27 | 3.38 | .000 |
| | | Native American | 63.57 | 6.56 | .000 |
| | | Pacific Islander | 46.96 | 7.19 | .000 |
| | | 2 or more races | 13.64 | 3.16 | .000 |
| | Black | Hispanic | -14.21 | 2.41 | .000 |
| | | Asian | -93.48 | 3.80 | .000 |
| | | Native American | 3.36 | 6.79 | 1.00 |
| | | Pacific Islander | -13.25 | 7.39 | 1.00 |
| | | 2 or more races | -46.57 | 3.60 | .000 |
| | Hispanic | Asian | -78.97 | 3.51 | .000 |
| | | Native American | 17.88 | 6.63 | .147 |
| | | Pacific Islander | 1.26 | 7.25 | 1.00 |
| | | 2 or more races | -32.05 | 3.30 | .000 |
| | Asian | Native American | 96.84 | 7.25 | .000 |
| | | Pacific Islander | 80.23 | 7.82 | .000 |
| | | 2 or more races | 46.91 | 4.41 | .000 |
| | Native American | Pacific Islander | -16.61 | 9.63 | 1.00 |
| | | 2 or more races | -49.93 | 7.15 | .000 |
| | Pacific Islander | 2 or more races | 33.32 | 7.73 | .000 |
| Math Plausible Value #2 | White | Black | 60.25 | 2.24 | .000 |
| | | Hispanic | 45.76 | 1.69 | .000 |
| | | Asian | -35.29 | 3.41 | .000 |
| | | Native American | 69.21 | 6.62 | .000 |
| | | Pacific Islander | 48.19 | 7.26 | .000 |
| | | 2 or more races | 15.05 | 3.19 | .000 |
| | Black | Hispanic | -14.49 | 2.43 | .000 |
| | | Asian | -95.53 | 2.44 | .000 |
| | | Native American | 8.97 | 6.85 | 1.00 |
| | | Pacific Islander | -12.06 | 7.47 | 1.00 |

| | | | | | |
|---|---|---|---|---|---|
| | | 2 or more races | -45.20 | 3.64 | .000 |
| | Hispanic | Asian | -81.05 | 3.54 | .000 |
| | | Native American | 23.45 | 6.69 | .010 |
| | | Pacific Islander | 2.42 | 7.32 | 1.00 |
| | | 2 or more races | -30.71 | 3.33 | .000 |
| | Asian | Native American | 104.50 | 7.32 | .000 |
| | | Pacific Islander | -21.03 | 9.72 | .000 |
| | | 2 or more races | 50.33 | 4.46 | .000 |
| | Native American | Pacific Islander | -21.03 | 9.73 | .644 |
| | | 2 or more races | -54.17 | 7.22 | .000 |
| | Pacific Islander | 2 or more races | -33.14 | 7.80 | .000 |
| Math Plausible Value #3 | White | Black | 61.74 | 2.24 | .000 |
| | | Hispanic | 45.96 | 1.69 | .000 |
| | | Asian | -33.92 | 3.41 | .000 |
| | | Native American | 66.86 | 6.63 | .000 |
| | | Pacific Islander | 43.34 | 7.26 | .000 |
| | | 2 or more races | 13.60 | 3.19 | .000 |
| | Black | Hispanic | -15.78 | 2.43 | .000 |
| | | Asian | -95.66 | 3.83 | .000 |
| | | Native American | 5.12 | 6.85 | 1.00 |
| | | Pacific Islander | -18.40 | 7.46 | .288 |
| | | 2 or more races | -48.14 | 3.64 | .000 |
| | Hispanic | Asian | -79.88 | 3.54 | .000 |
| | | Native American | 20.90 | 6.69 | .038 |
| | | Pacific Islander | -2.62 | 7.32 | 1.00 |
| | | 2 or more races | -32.36 | 3.33 | .000 |
| | Asian | Native American | 100.78 | 7.32 | .000 |
| | | Pacific Islander | 77.26 | 7.90 | .000 |
| | | 2 or more races | 47.12 | 4.56 | .000 |
| | Native American | Pacific Islander | -23.52 | 9.72 | .328 |
| | | 2 or more races | -53.26 | 7.22 | .000 |
| | Pacific Islander | 2 or more races | -29.74 | 7.80 | .003 |
| Math Plausible Value #4 | White | Black | 60.54 | 2.23 | .000 |
| | | Hispanic | 45.73 | 1.69 | .000 |
| | | Asian | -34.88 | 3.40 | .000 |
| | | Native American | 62.95 | 6.60 | .000 |
| | | Pacific Islander | 39.77 | 7.23 | .000 |

| | | | | | |
|---|---|---|---|---|---|
| | | 2 or more races | 17.51 | 3.18 | .000 |
| | Black | Hispanic | 14.81 | 2.43 | .000 |
| | | Asian | -95.41 | 3.82 | .000 |
| | | Native American | 2.41 | 6.83 | 1.00 |
| | | Pacific Islander | -20.77 | 7.44 | .110 |
| | | 2 or more races | -43.03 | 3.62 | .000 |
| | Hispanic | Asian | -80.61 | 3.53 | .000 |
| | | Native American | 17.21 | 6.69 | .207 |
| | | Pacific Islander | -5.97 | 7.29 | 1.00 |
| | | 2 or more races | -28.22 | 3.32 | .000 |
| | Asian | Native American | 97.82 | 7.29 | .000 |
| | | Pacific Islander | 74.64 | 7.87 | .000 |
| | | 2 or more races | 52.39 | 4.44 | .000 |
| | Native American | Pacific Islander | -23.18 | 9.69 | .352 |
| | | 2 or more races | -45.43 | 7.19 | .000 |
| | Pacific Islander | 2 or more races | -22.25 | 7.77 | .088 |
| Math Plausible Value #5 | White | Black | 59.32 | 2.24 | .000 |
| | | Hispanic | 45.50 | 1.69 | .000 |
| | | Asian | -36.04 | 3.41 | .000 |
| | | Native American | 64.64 | 6.63 | .000 |
| | | Pacific Islander | 44.57 | 7.26 | .000 |
| | | 2 or more races | 14.78 | 3.19 | .000 |
| | Black | Hispanic | -13.81 | 2.43 | .000 |
| | | Asian | -95.36 | 3.84 | .000 |
| | | Native American | 5.32 | 6.85 | 1.00 |
| | | Pacific Islander | -14.75 | 7.47 | 1.00 |
| | | 2 or more races | -44.54 | 3.64 | .000 |
| | Hispanic | Asian | 81.55 | 3.54 | .000 |
| | | Native American | 19.14 | 6.69 | .090 |
| | | Pacific Islander | -.93 | 7.32 | 1.00 |
| | | 2 or more races | -30.73 | 3.33 | .000 |
| | Asian | Native American | 100.68 | 7.32 | .000 |
| | | Pacific Islander | 80.61 | 7.90 | .000 |
| | | 2 or more races | 50.82 | 4.46 | .000 |
| | Native American | Pacific Islander | -20.07 | 9.73 | .821 |
| | | 2 or more races | -49.86 | 7.22 | .000 |
| | Pacific Islander | 2 or more races | -29.79 | 7.80 | .003 |

on two of the five Plausible Values developed. These significant differences arose on Plausible Value #2, $F(1, 10350) = 4.90$; $p = .037$, and on Plausible Value #5, $F(1, 10350) = 10.00$; $p = .002$. Male students did achieve higher mean scores than female students on the other three Plausible Values, as well, but the differences between the scores did not reach a level of statistical significance ($p < .05$).

While the developers of the TIMSS 2011 created five different Plausible Values for the Math Assessment to account for the possible differences in students' outcomes, there is also reason to believe that, with such a large sample size, well researched and developed scales, and near-random sampling procedure, the outcomes across each of the Plausible Values should remain consistent. Therefore, for the purposes of clarity and succinctness, not all five Plausible Values will be included in each analysis moving forward. However, given the demonstrated variability in the gender differences of achievement across the math assessment Plausible Values, it is necessary to use more than one Math Plausible Value. Therefore, Math Plausible Values #3 and #5 will be used in all subsequent analysis as these two values represent the least amount of gender differences and the greatest amount of gender differences, respectively.

When students' differences in math performance related to their racial demographic affiliation are considered, there are found to be statistically significant differences in the level of performance across all five Plausible Values (Plausible Value #1 = $F(6, 10340) = 252.5$; $p < .001$; Plausible Value #2 = $F(6, 10340) = 253.5$; $p < .001$; Plausible Value #3 = $F(6, 10340) = 255.7$; $p < .001$; Plausible Value #4 = $F(6, 10340) = 251.7$; $p < .001$; Plausible Value #5 = $F(6, 10340) = 248.2$; $p < .001$). In order to examine the pairwise comparisons between the racial groups, additional post-hoc

Bonferroni-adjusted analyses were completed. These analyses indicated that Asian students scored significantly higher on the math assessment than all other racial demographic groups across all five Plausible Values. Next, the analyses showed that White students scored significantly higher than all other groups of students, except for Asian students, across all five Plausible Values. Students who identified as Multi-Racial were found to earn significantly higher scores than Black, Hispanic, and Native American students across all five Plausible Values, and significantly outperformed Pacific Islander students on four of the five Plausible Values. Finally, Hispanic students were shown to earn significantly higher scores than Black students on all five Plausible Values, and significantly outperformed Native American students on two of the five Plausible Values. Due to the consistent findings of significantly different levels of performance according to racial demographic category, only one Plausible Value will be used in subsequent analyses. Plausible Value #1 will be used as it represents the median measured race-based difference in math achievement.

These scores indicate that the intended covariate of math performance fails to be independent of the racial demographic treatment variable. However, math performance was found to be independent of the gender treatment variable when Plausible Value #3 is used, but the two variables are not independent when Plausible Value #5 is used. In the subsequent ANCOVA analysis, calculations will be completed twice, one time using Plausible Value #3 as the covariate and once using Plausible Value #5 as the covariate. By examining the amount of variance explained after using each covariate, the impact of this independence assumption can be monitored.

**Differences in scores on the Science Assessment Plausible Values.** Similarly, gender and racial demographic differences on the Science Assessment Plausible Values were examined. Additional ANOVA analyses were completed to determine if those Plausible Values differed between groups. The same interpretative guidelines used with the Mathematics Plausible Values will be used with the Science Plausible Values as well. The ANOVA results between demographic groups in terms of science abilities are presented in Table 13, while Table 14 presents the Bonferroni-adjusted pairwise comparisons between racial demographic groups.

Table 13

*ANOVA Results for Comparisons between Demographic Categories and Science Assessment Scores*

| Scale | | Sum of Squares | Df | Mean Square | F | Sig |
|---|---|---|---|---|---|---|
| | | | Gender | | | |
| Plausible Value #1 | Between Groups | 303910.5 | 1 | 303910.5 | 46.66 | .000 |
| | Within Groups | 67387612.1 | 10350 | 6513.4 | | |
| | Total | 67691522.6 | 10350 | | | |
| Plausible Value #2 | Between Groups | 355356.4 | 1 | 355356.4 | 55.80 | .000 |
| | Within Groups | 65887266.9 | 10350 | 6368.4 | | |
| | Total | 66242623.2 | 10350 | | | |
| Plausible Value #3 | Between Groups | 356928.7 | 1 | 356928.7 | 56.79 | .000 |
| | Within Groups | 65029151.6 | 10350 | 6285.4 | | |
| | Total | 65386080.3 | 10350 | | | |
| Plausible Value #4* | Between Groups | 336338.9 | 1 | 336338.9 | 52.59 | .000 |
| | Within Groups | 66163521.6 | 10350 | 6295.1 | | |
| | Total | 66499860.5 | 10350 | | | |
| Plausible Value #5 | Between Groups | 330171.3 | 1 | 330171.3 | 51.18 | .000 |
| | Within Groups | 66739315.4 | 10350 | 6350.7 | | |
| | Total | 67069486.7 | 10350 | | | |
| | | | Race | | | |
| Plausible* Value #1 | Between Groups | 10737368.5 | 6 | 1789561.4 | 324.9 | .000 |
| | Within Groups | 56954154.1 | 10340 | 5406.4 | | |
| | Total | 67691522.6 | 10350 | | | |
| Plausible Value #2 | Between Groups | 10335129.8 | 6 | 1722521.6 | 318.6 | .000 |
| | Within Groups | 55907493.5 | 10340 | 5406.4 | | |
| | Total | 66424623.3 | 10350 | | | |
| Plausible Value #3 | Between Groups | 10415319.4 | 6 | 1735886.6 | 326.6 | .000 |
| | Within Groups | 54970761.0 | 10340 | 5315.8 | | |
| | Total | 65386080.4 | 10350 | | | |
| Plausible Value #4 | Between Groups | 10480004.2 | 6 | 1746667.4 | 322.4 | .000 |
| | Within Groups | 56019856.2 | 10340 | 5417.3 | | |
| | Total | 66499860.5 | 10350 | | | |
| Plausible Value #5 | Between Groups | 10780519.3 | 6 | 1796753.2 | 330.1 | .000 |
| | Within Groups | 56288967.4 | 10340 | 5443.3 | | |
| | Total | 67069486.7 | 10350 | | | |

Note: * denotes Plausible Value that will be used during subsequent analyses

Table 14

*Mean Difference Results of Bonferroni-Adjusted Pairwise Post-Hoc Analyses between Students' Race and TIMSS 2011 Science Assessment Plausible Values*

| Scale | (I) Race | (J) Race | Mean Difference (I-J) | Std. Error | Sig. |
|---|---|---|---|---|---|
| Science Plausible Value #1 | White | Black | 77.42 | 2.34 | .000 |
| | | Hispanic | 60.29 | 1.77 | .000 |
| | | Asian | -.16 | 3.56 | 1.00 |
| | | Native American | 66.40 | 6.90 | .000 |
| | | Pacific Islander | 76.34 | 7.57 | .000 |
| | | 2 or more races | 18.33 | 3.33 | .000 |
| | Black | Hispanic | -17.14 | 2.54 | .000 |
| | | Asian | -77.58 | 4.00 | .000 |
| | | Native American | -11.03 | 7.15 | 1.00 |
| | | Pacific Islander | -1.09 | 7.78 | 1.00 |
| | | 2 or more races | -59.09 | 3.79 | .000 |
| | Hispanic | Asian | -60.45 | 1.77 | .000 |
| | | Native American | 6.11 | 6.98 | 1.00 |
| | | Pacific Islander | 16.05 | 7.63 | .745 |
| | | 2 or more races | -41.95 | 3.47 | .000 |
| | Asian | Native American | 66.55 | 7.63 | .000 |
| | | Pacific Islander | 76.50 | 8.23 | .000 |
| | | 2 or more races | 18.49 | 4.65 | .001 |
| | Native American | Pacific Islander | 9.94 | 10.14 | 1.00 |
| | | 2 or more races | -48.06 | 7.53 | .000 |
| | Pacific Islander | 2 or more races | -58.00 | 8.14 | .000 |
| Science Plausible Value #2 | White | Black | 77.21 | 2.32 | .000 |
| | | Hispanic | 58.58 | 1.75 | .000 |
| | | Asian | 2.40 | 3.53 | 1.00 |
| | | Native American | 68.48 | 6.85 | .000 |
| | | Pacific Islander | 71.37 | 7.50 | .000 |
| | | 2 or more races | 17.54 | 2.30 | .000 |
| | Black | Hispanic | -18.63 | 2.52 | .000 |
| | | Asian | -74.82 | 3.96 | .000 |
| | | Native American | -8.73 | 7.08 | 1.00 |
| | | Pacific Islander | -5.84 | 7.71 | 1.00 |

| | | | | | |
|---|---|---|---|---|---|
| | | 2 or more races | -59.67 | 3.76 | .000 |
| | Hispanic | Asian | -56.19 | 3.66 | .000 |
| | | Native American | 9.90 | 6.92 | 1.00 |
| | | Pacific Islander | 12.79 | 7.56 | 1.00 |
| | | 2 or more races | -41.04 | 3.44 | .000 |
| | Asian | Native American | 66.08 | 7.56 | .000 |
| | | Pacific Islander | 68.97 | 8.16 | .000 |
| | | 2 or more races | 15.14 | 4.60 | .021 |
| | Native American | Pacific Islander | 2.89 | 10.05 | 1.00 |
| | | 2 or more races | -50.94 | 7.46 | .000 |
| | Pacific Islander | 2 or more races | -53.83 | 8.06 | .000 |
| Science Plausible Value #3 | White | Black | 77.49 | 2.30 | .000 |
| | | Hispanic | 58.56 | 1.73 | .000 |
| | | Asian | .08 | 3.50 | 1.00 |
| | | Native American | 67.59 | 6.79 | .000 |
| | | Pacific Islander | 73.63 | 7.43 | .000 |
| | | 2 or more races | 17.67 | 3.27 | .000 |
| | Black | Hispanic | -19.03 | 2.50 | .000 |
| | | Asian | -77.41 | 3.93 | .000 |
| | | Native American | -9.93 | .702 | 1.00 |
| | | Pacific Islander | -3.86 | 7.65 | 1.00 |
| | | 2 or more races | -59.81 | 3.73 | .000 |
| | Hispanic | Asian | -58.38 | 3.63 | .000 |
| | | Native American | 9.10 | 6.86 | 1.00 |
| | | Pacific Islander | 15.17 | 7.50 | .903 |
| | | 2 or more races | -40.78 | 3.41 | .000 |
| | Asian | Native American | 67.48 | 7.50 | .000 |
| | | Pacific Islander | 73.56 | 8.09 | .000 |
| | | 2 or more races | 17.60 | 4.56 | .002 |
| | Native American | Pacific Islander | 6.07 | 9.96 | 1.00 |
| | | 2 or more races | -49.88 | 7.40 | .000 |
| | Pacific Islander | 2 or more races | -55.96 | 7.99 | .000 |
| Science Plausible Value #4 | White | Black | 76.58 | 2.31 | .000 |
| | | Hispanic | 59.08 | 1.75 | .000 |
| | | Asian | -1.63 | 3.52 | 1.00 |
| | | Native American | 70.15 | 6.85 | .000 |
| | | Pacific Islander | 72.42 | 7.51 | .000 |

| | | | | | |
|---|---|---|---|---|---|
| | | 2 or more races | 16.49 | 3.30 | .000 |
| | Black | Hispanic | -17.51 | 2.52 | .000 |
| | | Asian | -78.21 | 3.97 | .000 |
| | | Native American | -6.43 | 7.09 | 1.00 |
| | | Pacific Islander | -4.16 | 7.72 | 1.00 |
| | | 2 or more races | -60.09 | 3.76 | .000 |
| | Hispanic | Asian | -60.70 | 3.66 | .000 |
| | | Native American | 11.08 | 6.92 | 1.00 |
| | | Pacific Islander | 13.34 | 7.57 | 1.00 |
| | | 2 or more races | -42.58 | 3.44 | .000 |
| | Asian | Native American | 71.78 | 7.57 | .000 |
| | | Pacific Islander | 74.05 | 8.17 | .000 |
| | | 2 or more races | 18.12 | 4.61 | .002 |
| | Native American | Pacific Islander | 2.27 | 10.06 | 1.00 |
| | | 2 or more races | -53.66 | 7.47 | .000 |
| | Pacific Islander | 2 or more races | -55.93 | 8.07 | .000 |
| Science Plausible Value #5 | White | Black | 77.96 | 2.32 | .000 |
| | | Hispanic | 60.18 | 1.76 | .000 |
| | | Asian | 2.03 | 3.54 | 1.00 |
| | | Native American | 68.93 | 6.87 | .000 |
| | | Pacific Islander | 77.69 | 7.52 | .000 |
| | | 2 or more races | 17.12 | 3.31 | .000 |
| | Black | Hispanic | -17.78 | 2.32 | .000 |
| | | Asian | -75.93 | 3.98 | .000 |
| | | Native American | -9.03 | 7.10 | 1.00 |
| | | Pacific Islander | -.27 | 7.74 | 1.00 |
| | | 2 or more races | -60.84 | 3.77 | .000 |
| | Hispanic | Asian | -58.15 | 3.67 | .000 |
| | | Native American | 8.75 | 6.94 | 1.00 |
| | | Pacific Islander | 17.51 | 7.59 | .442 |
| | | 2 or more races | -43.06 | 3.45 | .000 |
| | Asian | Native American | 66.90 | 7.59 | .000 |
| | | Pacific Islander | 75.66 | 8.19 | .000 |
| | | 2 or more races | 15.09 | 4.62 | .023 |
| | Native American | Pacific Islander | 8.76 | 10.08 | 1.00 |
| | | 2 or more races | -51.81 | 7.48 | .000 |
| | Pacific Islander | 2 or more races | -60.57 | 8.09 | .000 |

Results on the TIMSS 2011 science assessment show a similar pattern of results in terms of gender and racial differences. Across each of the five Plausible Values calculated, male students significantly outperformed female students in terms of science achievement (Plausible Value #1 = $F(1, 10350) = 46.66$; $p < .001$; Plausible Value #2 = $F(1, 10350) = 55.80$; $p < .001$; Plausible Value #3 = $F(1, 10350) = 56.79$; $p < .001$; Plausible Value #4 = $F(1, 10350) = 52.59$; $p < .001$; Plausible Value #5 = $F(1, 10350) = 51.18$; $p < .001$). Due to the consistent findings of significantly different levels of science performance according to students' reported gender, only one Plausible Value will be used in subsequent analyses. Plausible Value #4 will be used as it represents the median measured gender-based difference in science achievement.

When examining the relationship between students' race and performance on the Science Assessment, results were also consistent across each of the five Plausible Values. Students' performance across each Plausible Value was found to vary significantly according to the students' reports of their racial group affiliation (Plausible Value #1 = $F(6, 10340) = 324.9$; $p < .001$; Plausible Value #2 = $F(6, 10340) = 318.6$; $p < .001$; Plausible Value #3 = $F(6, 10340) = 326.6$; $p < .001$; Plausible Value #4 = $F(6, 10340) = 322.4$; $p < .001$; Plausible Value #5 = $F(6, 10340) = 330.1$; $p < .001$). Further analyses using the Bonferroni-adjustment to examine the pairwise differences between each racial demographic groups' level of performance also indicated consistent results. Across each of the five Plausible Values, White and Asian students significantly outperformed all other racial groups of students. White and Asian students, however, were not found to perform at a significantly different level across any of the Science Assessment values. Students who reported belonging to two or more racial demographic groups were shown to significantly outperform Black, Hispanic, Native American, and Pacific Islander students. Finally, Hispanic students significantly outperformed Black students across each of the five Plausible Values. Due to the consistency of these results, only Science Assessment Plausible Value #1 will be used during the subsequent analyses in which science achievement and race are used as variables.

Similar to the findings reported regarding the Math Assessment, the Science Assessment covariate does not appear to be independent of either the Gender or Race independent variables. As discussed above, this lack of independence will complicate the interpretation of the subsequent ANCOVA analysis as the independent variable and covariates will account for a shared portion of the explained variance in the resulting model.

## Homogeneity of Regression Slopes

In order to measure if the relationship between the Student Bullied at School Scale and the covariates are consistent across each of the treatment groups, the homogeneity of the various regression slopes was calculated. This is meant to ensure that each treatment group experiences similar effects due to each covariate. These relationships between the covariates were calculated for both the Gender and Race treatment variables. The results are presented below.

**Homogeneity of regression slopes between covariates and Gender.** The first set of covariates measured were those covariates used to assess students' perceptions and performance in math. Therefore, the covariates of performance on the Math Assessment, and students' responses to the Student Likes Learning Math Scale and Student Confident in Math Scale were used. The effect of the interaction between these covariate scales,

Table 15

*Tests Measuring the Homogeneity of Regression Slopes between Math-Based Covariates, Using Math Assessment Plausible Value #3, and the Gender Treatment Variable on the Student Bullied at School Scale Dependent Variable*

| Covariate Scale | Sum of Squares | Df | Mean Square | F | Sig |
|---|---|---|---|---|---|
| Gender | .05 | 1 | .05 | .01 | .912 |
| Math Assessment Plausible Value #3 | 27.86 | 1 | 27.86 | 7.57 | .006 |
| Student Likes Learning Math Scl | 29.98 | 1 | 29.98 | 8.14 | .004 |
| Student Confident in Math Scl | 81.59 | 1 | 81.59 | 22.16 | .000 |
| Gender * Math Assessment PV #3 | 3.39 | 1 | 3.39 | .92 | .338 |
| Gender * Student Likes Learning Math | .22 | 1 | .22 | .06 | .809 |
| Gender * Student Confident in Math | 1.99 | 1 | 1.99 | .54 | .463 |
| Gender * Math Assessment * Like Learning Math * Confident in Math | 34.78 | 2 | 17.39 | 4.72 | .009 |
| Error | 38066.15 | 10340 | 3.68 | | |
| Corrected Total | 38288.85 | 10350 | | | |

using Math Assessment Plausible Value #3, and students' Gender on the Student Bullied at School outcome variable is reported in Table 15, while the covariate effect using Math Assessment Plausible Value #5 is reported in Table 16.

Results indicate that the pairwise interactions between the regression slopes of the Gender treatment variable and each of the three covariates are not significant. This is true when either Plausible Value #3 (Gender × Math Assessment PV #3 = $F(1, 10340)$ =

Table 16

*Tests Measuring the Homogeneity of Regression Slopes between Math-Based Covariates, Using Math Assessment Plausible Value #5, and the Gender Treatment Variable on the Student Bullied at School Scale Dependent Variable*

| Covariate Scale | Sum of Squares | Df | Mean Square | F | Sig |
|---|---|---|---|---|---|
| Gender | .38 | 1 | .38 | .10 | .748 |
| Math Assessment Plausible Value #5 | 36.15 | 1 | 36.15 | 9.82 | .002 |
| Student Likes Learning Math Scl | 35.38 | 1 | 35.38 | 9.61 | .002 |
| Student Confident in Math Scl | 88.56 | 1 | 88.56 | 24.06 | .000 |
| Gender * Math Assessment PV #5 | 5.57 | 1 | 5.57 | 1.51 | .219 |
| Gender * Student Likes Learning Math | .03 | 1 | .03 | .01 | .928 |
| Gender * Student Confident in Math | 1.44 | 1 | 1.44 | .39 | .531 |
| Gender * Math Assessment * Likes Learning Math * Confident in Math | 40.72 | 2 | 20.36 | 5.53 | .004 |
| Error | 38054.62 | 10340 | 3.68 | | |
| Corrected Total | 38288.85 | 10350 | | | |

.92; $p$ = .338; Gender × Student Likes Learning Math = $F(1, 10340)$ = .06; $p$ = .809; Gender × Student Confident Learning Math = $F(1, 10340)$ = .54; $p$ = .463) or Plausible Value #5 (Gender × Math Assessment PV #5 = $F(1, 10340)$ = 1.51; $p$ = .219; Gender × Student Likes Learning Math = $F(1, 10340)$ = .01; $p$ = .928; Gender × Student Confident Learning Math = $F(1, 10340)$ = .39; $p$ = .531) are used. This suggests that the regression slopes for each individual covariate do not differ significantly depending on the respondent's gender.

However, when the interaction between each of the three covariates together and the treatment variable of gender is examined, the regression slopes are found to be significantly different when using both Math Assessment Plausible Value #3 (Gender × Math Assessment PV #3 × Student Likes Learning Math × Student Confident Learning Math = $F(2, 10340) = 4.72$; $p = .009$) and Math Assessment Plausible Value #5 (Gender × Math Assessment PV #5 × Student Likes Learning Math × Student Confident Learning Math = $F(2, 10340) = 5.53$; $p = .004$). This indicates that when each of the math covariates are added into the final ANCOVA procedure, the resulting regression slopes are significantly different, indicating that the covariates do not vary consistently across the gender groups. Therefore, the Math data does violate the assumption of homogeneity of regression slopes when Gender is used as the treatment variable. This further complicates the interpretation of the final ANCOVA analysis.

The next set of covariates measured were those covariates used to assess students' perceptions and performance in science. Therefore, the covariates of performance on the Science Assessment, and students' responses to the Student Likes Learning Science Scale and Student Confident in Science Scale were used. The effect of the interaction between these covariate scales, using Science Assessment Plausible Value #4, and students' Gender on the Student Bullied at School outcome variable is reported in Table 17.

Results indicate that, again, the pairwise interactions between the regression slopes of the Gender treatment variable and each of the three covariates are not significant (Gender × Science Assessment PV #3 = $F(1, 10340) = 2.17$; $p = .141$; Gender

Table 17

*Tests Measuring the Homogeneity of Regression Slopes between Science-Based Covariates, Using Science Assessment Plausible Value #4, and the Gender Treatment Variable on the Student Bullied at School Scale Dependent Variable*

| Covariate Scale | Sum of Squares | Df | Mean Square | F | Sig |
|---|---|---|---|---|---|
| Gender | .63 | 1 | .63 | .17 | .679 |
| Science Assessment Plausible Value #4 | .61 | 1 | .61 | .17 | .685 |
| Student Likes Learning Science Scl | 1.00 | 1 | 1.00 | .270 | .603 |
| Student Confident in Science Scl | 60.07 | 1 | 60.07 | 18.76 | .000 |
| Gender * Science Assessment PV #4 | 7.98 | 1 | 7.98 | 2.17 | .141 |
| Gender * Student Likes Learning Science | .52 | 1 | .52 | .14 | .708 |
| Gender * Student Confident in Science | 2.95 | 1 | 2.95 | .80 | .371 |
| Gender * Science Assessment * Like Learning Science * Confident in Science | 8.86 | 2 | 4.43 | 1.20 | .300 |
| Error | 38061.32 | 10340 | 3.68 | | |
| Corrected Total | 38288.85 | 10350 | | | |

× Student Likes Learning Science = $F(1, 10340) = .14$; $p = .708$; Gender × Student Confident Learning Science = $F(1, 10340) = .80$; $p = .371$). This suggests that the regression slopes for each individual covariate do not differ significantly depending on the respondent's gender. The science covariates also demonstrate a non-significant interaction between each of the three covariates together and the treatment variable of gender (Gender × Science Assessment PV #4 × Student Likes Learning Science ×

Student Confident in Science = $F(2, 10340) = 1.20$; $p = .300$). This indicates that when each of the science covariates are added into the final ANCOVA procedure, the resulting regression slopes are not significantly different and the covariates vary consistently across the gender groups. Therefore, the Science data does not violate the assumption of homogeneity of regression slopes when Gender is used as the treatment variable.

**Homogeneity of regression slopes between covariates and Race.** The next set of covariates measured were those covariates used to assess students' perceptions and performance in math, and how those covariates relate to the Race treatment variable. Similar to above, the covariates of performance on the Math Assessment, and students' responses to the Student Likes Learning Math Scale and Student Confident in Math Scale were used. The effect of the interaction between these covariate scales, using Math Assessment Plausible Value #1, and students' Race on the Student Bullied at School outcome variable is reported in Table 18.

Results indicate that the pairwise interactions between the regression slopes of the Race treatment variable and both the Student Likes Learning Math Scale and the Student Confident in Math Scale covariates are significant (Race × Student Likes Learning Math = $F(6, 10310) = 2.74$; $p = .012$; Race × Student Confident in Math = $F(6, 10310) = 2.76$; $p = .011$). However, the interaction between the Race treatment variable and the Math Assessment Plausible Value #1 covariate was not significant (Gender × Math Assessment PV #1 = $F(1, 10310) = 1.11$; $p = .352$).

Consistent with the results when using the Gender independent variable, when the interaction between each of the three covariates together and the treatment variable of Race is examined, the regression slopes are found to be significantly different (Race ×

Table 18

*Tests Measuring the Homogeneity of Regression Slopes between Math-Based Covariates, Using Math Assessment Plausible Value #1, and the Race Treatment Variable on the Student Bullied at School Scale Dependent Variable*

| Covariate Scale | Sum of Squares | Df | Mean Square | F | Sig |
|---|---|---|---|---|---|
| Race | 46.78 | 6 | 7.80 | 2.12 | .047 |
| Math Assessment Plausible Value #1 | 8.03 | 1 | 8.03 | 2.19 | .139 |
| Student Likes Learning Math Scl | 3.59 | 1 | 3.59 | .98 | .323 |
| Student Confident in Math Scl | 14.84 | 1 | 14.84 | 4.05 | .044 |
| Race * Math Assessment PV #1 | 24.44 | 6 | 4.07 | 1.11 | .352 |
| Race * Student Likes Learning Math | 60.29 | 6 | 10.05 | 2.74 | .012 |
| Race * Student Confident in Math | 60.62 | 6 | 10.10 | 2.76 | .011 |
| Race * Math Assessment * Like Learning Math * Confident in Math | 107.79 | 7 | 15.40 | 4.20 | .000 |
| Error | 37790.84 | 10310 | 3.66 | | |
| Corrected Total | 38288.85 | 10350 | | | |

Math Assessment PV #1 × Student Likes Learning Math × Student Confident in Math = $F(7, 10310) = 4.20$; $p < .001$). This indicates that when each of the math covariates are added into the final ANCOVA procedure, the resulting regression slopes are significantly different and signifies that the covariates do not vary consistently across the racial demographic groups. Therefore, the Math-related data does violate the assumption of

Table 19

*Tests Measuring the Homogeneity of Regression Slopes between Science-Based Covariates, Using Science Assessment Plausible Value #1, and the Race Treatment Variable on the Student Bullied at School Scale Dependent Variable*

| Covariate Scale | Sum of Squares | Df | Mean Square | F | Sig |
|---|---|---|---|---|---|
| Race | 51.02 | 6 | 8.50 | 2.32 | .031 |
| Science Assessment Plausible Value #1 | 5.39 | 1 | 5.39 | 1.47 | .226 |
| Student Likes Learning Science Scl | .76 | 1 | .76 | .21 | .649 |
| Student Confident in Science Scl | 6.15 | 1 | 6.15 | 1.68 | .196 |
| Race * Science Assessment PV #1 | 42.04 | 6 | 7.01 | 1.91 | .075 |
| Race * Student Likes Learning Science | 46.03 | 6 | 7.67 | 2.09 | .051 |
| Race * Student Confident in Science | 36.37 | 6 | 6.06 | 1.65 | .129 |
| Race * Science Assessment * Like Learning Science * Confident in Science | 52.85 | 7 | 7.55 | 2.06 | .045 |
| Error | 37852.54 | 10310 | 3.67 | | |
| Corrected Total | 38288.85 | 10350 | | | |

homogeneity of regression slopes when Race is used as the treatment variable. This further complicates the interpretation of the final ANCOVA analysis.

The last set of covariates measured were those covariates used to assess students' perceptions and performance in science. Therefore, the covariates of performance on the Science Assessment, and students' responses to the Student Likes Learning Science Scale and Student Confident in Science Scale were used. The effect of the interaction between these covariate scales, using Science Assessment Plausible Value #1, and students' Race on the Student Bullied at School Scale outcome variable is reported in Table 19.

Results indicate that, when analyzing the pairwise interactions between the regression slopes of the Race treatment variable and each of the three covariate scales, the relationships do not reach a level of statistical significance (Race × Science Assessment PV #1 = $F(6, 10310) = 1.91$; $p = .075$; Race × Student Likes Learning Science = $F(6, 10310) = 2.09$; $p = .051$; Race × Student Confident in Science = $F(6, 10310) = 1.65$; $p = .129$). This suggests that the regression slopes for each individual covariate do not differ significantly depending on the respondent's race. The science covariates do, however, demonstrate a significant interaction when each of the three covariate scales are considered together and compared to the Race treatment variable (Race × Science Assessment PV #1 × Student Likes Learning Science × Student Confident in Science = $F(7, 10310) = 2.06$; $p = .045$). This indicates that when analyzing the science-related data using the Race treatment variable, the data violate the assumption of homogeneity of regression slopes when Race is used as the treatment variable.

## Preliminary Analyses

### Correlations between Frequency of Bullying Victimization and Academic Performance

To gather more information regarding the relationship between a student's responses on the Student Bullied at School Scale and his or her performance on the Math and Science Assessments, Pearson correlations were calculated between these variables. The resulting correlations gathered from the overall student sample are presented in Table 20. The correlations between reports of the frequency of bullying victimization and

Table 20

*Overall Pearson Correlations between Students' Math and Science Performance on the Skills Assessment Plausible Values and Those Students' Reports on the Student Bullied at School Scale*

| | Math Assessment Plausible Value #1 | Math Assessment Plausible Value #3 | Math Assessment Plausible Value #5 | Science Assessment Plausible Value #1 | Science Assessment Plausible Value #4 |
|---|---|---|---|---|---|
| Student Bullied at School Scale | .040** | .029** | .031** | -.012 | -.014 |
| Math Assessment Plausible Value #1 | 1 | .914** | .915** | .833** | .803** |
| Math Assessment Plausible Value #3 | | 1 | .914** | .802** | .801** |
| Math Assessment Plausible Value #5 | | | 1 | .801** | .801** |
| Science Assessment Plausible Value #1 | | | | 1 | .887** |

Note: * denotes correlation is significant at $p < .05$ level
** denotes correlation is significant at $p < .01$ level

academic performance for each gender group are presented in Table 21, and such correlations for each racial demographic group are presented in Table 22.

These correlations indicate that, when the entire student sample is considered together, students' responses on the Student Bullied at School Scale are significantly and positively related to their performance on the Math Assessments as measured by Plausible Value #1, $r = .040$; $p < .01$, Plausible Value #3, $r = .029$; $p < .01$, and Plausible Value #5, $r = .031$; $p < .01$. When compared to their performance on the Science Assessments, however, students' responses to the Student Bullied at School scale were not significantly related to their science achievement, and the direction of this relationship was shown to be negative. This suggests that, in the overall sample, as

Table 21

*Pearson Correlations between Students' Math and Science Performance on the Skills Assessment Plausible Values and Those Students' Reports on the Student Bullied at School Scale Divided by Gender*

| | Math Assessment Plausible Value #3 | Math Assessment Plausible Value #5 | Science Assessment Plausible Value #4 |
|---|---|---|---|
| | Female Students ($N$ = 5,240) | | |
| Student Bullied at School Scale | .056** | .060** | .013 |
| Math Assessment Plausible Value #3 | 1 | .912** | .796** |
| Math Assessment Plausible Value #5 | | 1 | .796** |
| | Male Students ($N$ = 5,110) | | |
| Student Bullied at School Scale | .005 | .006 | -.035* |
| Math Assessment Plausible Value #3 | 1 | .916** | .807** |
| Math Assessment Plausible Value #5 | | 1 | .805** |

Note: * denotes correlation is significant at $p < .05$ level
** denotes correlation is significant at $p < .01$ level

Science Assessment Plausible Value scores increased, reports of bullying victimization frequency decreased. This was true for both Science Assessment Plausible Value #1, $r$ = -.012; $p$ = .22, and Plausible Value #4, $r$ = -.014; $p$ = .16.

However, correlational results appear to vary when data is divided my students' gender and race. When the data is divided by the Gender variable, female students are shown to demonstrate a significant positive relationship between their scores on the Student Bullied at School Scale and scores on the Math Assessment Plausible Value #3, $r$ = .056; $p < .01$, and Plausible Value #5, $r$ = .060; $p < .01$. However, this significant relationship disappears for male students. Correlations between bullying victimization

Table 22

*Pearson Correlations between Students' Math and Science Performance on the Skills Assessment Plausible Values and Those Students' Reports on the Student Bullied at School Scale Divided by Racial Demographic*

| | Math Assessment Plausible Value #1 | Science Assessment Plausible Value #1 |
|---|---|---|
| White Students (*N* = 5,170) | | |
| Student Bullied at School Scale | .061** | .013 |
| Math Assessment Plausible Value #1 | 1 | .805** |
| Black Students (*N* = 1,250) | | |
| Student Bullied at School Scale | .046 | -.001 |
| Math Assessment Plausible Value #1 | 1 | .824** |
| Hispanic Students (*N* = 2,680) | | |
| Student Bullied at School Scale | .048* | -.017 |
| Math Assessment Plausible Value #1 | 1 | .813** |
| Asian Students (*N* = 480) | | |
| Student Bullied at School Scale | .087 | .071 |
| Math Assessment Plausible Value #1 | 1 | .854** |
| Native American Students (*N* = 120) | | |
| Student Bullied at School Scale | -.073 | -.081 |
| Math Assessment Plausible Value #1 | 1 | .831** |
| Pacific Islander Students (*N* = 100) | | |
| Student Bullied at School Scale | .041 | .017 |
| Math Assessment Plausible Value #1 | 1 | .764** |
| Multi-Racial Students (*N* = 550) | | |
| Student Bullied at School Scale | .058 | -.015 |
| Math Assessment Plausible Value #1 | 1 | .799** |

Note: * denotes correlation is significant at $p < .05$ level
** denotes correlation is significant at $p < .01$ level

frequency and math performance were shown to be positive, but non-significant for both Math Assessment Plausible Value #3, $r = .005$; $p = .70$, and Plausible Value #5, $r = .006$; $p = .681$. This indicates that has math performance increases, female students are more likely to report a corresponding increase in the frequency of bullying victimization than are male students.

There continues to be a noticeable gender-based difference when science performance is examined. While the overall sample reported a negative relationship between science performance and frequency of bullying victimization, when results are divided by students' gender, male students' reports appear largely responsible for this negative relationship, $r = -.035$; $p = .01$. Data gathered from female students, however, indicate a positive, but non-significant, relationship between Science Assessment Plausible Value #4 and responses on the Student Bullied at School Scale, $r = .013$; $p = .34$. These patterns of results suggest that as male students earn higher scores on the Science Assessment, they report lower frequencies of bullying victimization. However, for female students, as they earn higher scores on the Science Assessment, they tend to report higher frequencies of bullying victimization.

When the correlations between Math and Science achievement and responses on the Student Bullied at School Scale are calculated for each racial demographic group, several differences also appear. For every racial group except Native American students, students' performance on the Math Assessment was positively correlated with reports of the frequency of bullying victimization. This indicates that for most students, as math performance increased, the frequency of bullying victimization tended to increase as well. However, only two groups of students, White, $r = .061$; $p < .01$, and Hispanic, $r = .048$; $p = .01$, demonstrated a correlation that reached a level of statistical significance. While the correlation between Asian students' math abilities and reports of bullying approached significance, $r = .087$; $p = .057$, the relationship did not reach the $p < .05$ level.

None of the correlations calculated between any of the racial groups' levels of achievement on the Science Assessment and their reports on the Student Bullied at School Scale reached a level of statistical significance. The direction of the relationships between these variables was inconsistent. Three racial demographic groups, White, Asian, and Pacific Islander students, reported slightly positive relationships, while the four other racial demographic groups, Black, Hispanic, Native American, and Multi-racial students, demonstrated slightly negative relationships. These results indicate that there is little to no difference in the relationship between math, and especially, science achievement and reports of bullying victimization based on students' racial affiliation.

**Direct Effects of Covariates on Dependent Variable**

In order to measure the direct effects that the covariates of Student Likes Learning Math/Science, Student Confident in Math/Science, and students' performance on the Math and Science Assessments have on the dependent variable of Students Bullied at School, Analysis of Variance (ANOVA) calculations were completed between each covariate and the dependent variable. Results of these ANOVA procedures are listed below.

Table 23

*ANOVA Investigation into the Independence between the Dependent Variable (Student Bullied at School Scale) and the Student Questionnaire Covariates*

| Covariate Scale | | Sum of Squares | Df | Mean Square | F | Sig |
|---|---|---|---|---|---|---|
| Students Like Learning Math Scale | Between Groups | 62.41 | 1 | 62.41 | 16.89 | .000 |
| | Within Groups | 38226.44 | 10350 | 3.70 | | |
| | Total | 38288.85 | 10350 | | | |
| Students Like Learning Science Scale | Between Groups | 17.15 | 1 | 17.15 | 4.63 | .031 |
| | Within Groups | 38271.70 | 10350 | 3.70 | | |
| | Total | 38288.85 | 10350 | | | |
| Students Confident in Math Scale | Between Groups | 135.05 | 1 | 135.05 | 36.62 | .000 |
| | Within Groups | 38153.81 | 10350 | 3.69 | | |
| | Total | 38288.85 | 10350 | | | |
| Students Confident in Science Scale | Between Groups | 103.15 | 1 | 103.15 | 27.97 | .000 |
| | Within Groups | 38185.70 | 10350 | 3.69 | | |
| | Total | 38288.85 | 10350 | | | |

**Student Questionnaire covariates.** The four covariates that were gathered from the Student Questionnaire are examined first. Table 23 presents the ANOVA results between the Students Bullied at School Scale and those covariates.

These results suggest that each of the covariates are significantly related to the dependent variable of the Students Bullied at School Scale. Given these significant relationships, more information was warranted to determine precisely how the covariates and the dependent variables were related. In order to obtain this information, another set of ANOVA analyses were completed. However, during these analyses, the Students Bullied at School Index was used as the dependent variable, rather than the Students Bullied at School Scale. As described above, the Index version of this variable developed cutpoints in the bullying victimization frequency data that comprises the Scale variable and created three independent groups referred to as “Almost Never,” “About Monthly,”

Table 24

*Demographic Data for Student Bullied at School Index by Student Questionnaire Covariates*

| Demographic Category (N) | Students Like Learning Math Scale | | Students Like Learning Science Scale | | Students Confident in Math Scale | | Students Confident in Science Scale | |
|---|---|---|---|---|---|---|---|---|
| | *M* | *SD* | *M* | *SD* | *M* | *SD* | *M* | *SD* |
| Almost Never (6,530) | 9.54 | 1.98 | 9.61 | 2.06 | 10.66 | 2.21 | 10.41 | 2.11 |
| About Monthly (2,880) | 9.47 | 2.10 | 9.58 | 2.01 | 10.42 | 2.29 | 10.24 | 2.10 |
| About Weekly (940) | 9.35 | 2.20 | 9.43 | 2.20 | 10.30 | 2.43 | 10.09 | 2.32 |
| Total (10,350) | 9.50 | 2.04 | 9.58 | 2.06 | 10.56 | 2.26 | 10.34 | 2.13 |

and "About Weekly." By using this Index variable, comparisons can be made between students who report low, medium, and high rates of bullying victimization in terms of their math and science confidence, enjoyment, and achievement. Table 24 presents the demographic breakdown for the Students Bullied at School Index, Table 25 presents the results of this ANOVA, while Table 26 presents the pairwise, post-hoc comparisons between the three bullying frequency groups.

After using the Index rather than the Scale version of the Student Bullied at School variable, significant differences remain between students' rates of reported bullying victimization and their reports of liking math, $F(2, 10350) = 4.41$; $p = .012$, liking science, $F(2, 10350) = 3.29$; $p = .037$, confidence in math, $F(2, 10350) = 17.59$; $p < .001$, and confidence in science, $F(2, 10350) = 13.61$; $p < .001$. Pairwise comparisons indicate that students whose responses fall into the "Almost Never" category of bullying frequency reported significantly higher rates of enjoyment in math and confidence in both

Table 25

*ANOVA Investigation into Differences in the Student Bullied at School Index and the Student Questionnaire Covariates*

| Covariate Scale | | Sum of Squares | Df | Mean Square | F | Sig |
|---|---|---|---|---|---|---|
| Student Likes Learning Math Scale | Between Groups | 36.57 | 2 | 18.28 | 4.41 | .012 |
| | Within Groups | 42912.18 | 10350 | 4.15 | | |
| | Total | 42948.75 | 10350 | | | |
| Student Likes Learning Science Scale | Between Groups | 27.89 | 2 | 13.95 | 3.29 | .037 |
| | Within Groups | 43902.11 | 10350 | 4.24 | | |
| | Total | 43930.00 | 10350 | | | |
| Student Confident in Math Scale | Between Groups | 179.58 | 2 | 89.79 | 17.59 | .000 |
| | Within Groups | 52795.51 | 10350 | 5.10 | | |
| | Total | 52975.09 | 10350 | | | |
| Student Confident in Science Scale | Between Groups | 123.23 | 2 | 61.62 | 13.61 | .000 |
| | Within Groups | 46832.98 | 10350 | 4.53 | | |
| | Total | 46956.20 | 10350 | | | |

math and science than students whose reports fall in the “Almost Weekly” category. Students in the “Almost Never” category also reported significantly higher rates of confidence in both math and science than students in the “Almost Monthly” bullying frequency category.

**Math and Science achievement covariates.** The relationship between students’ performance on the math and science assessments and the Student Bullied at School variable was tested next. Similar to the procedure used when testing the Student Questionnaire covariates, the relationships between the Math and Science Plausible Values was first tested against the Scale version of the dependent variable; then, in order to better understand how frequency of bullying victimization appears related to math and science achievement, an ANOVA analysis was completed between the Index version of the Student Bullied at School variable and the math and science achievement scores.

Table 26

*Mean Difference Results of Bonferroni-Adjusted Pairwise Post-Hoc Analyses between Scores on Student Bullied at School Index and Student Questionnaire Covariates*

| Covariate Scale | (I) Student Bullied at School Index Rating | (J) Student Bullied at School Index Rating | Mean Difference (I-J) | Std. Error | Sig. |
|---|---|---|---|---|---|
| Student Likes Learning Math Scale | Almost Never | About Monthly | .076 | .046 | .292 |
| | | About Weekly | .195 | .071 | .018 |
| | About Monthly | About Weekly | .119 | .077 | .358 |
| Student Likes Learning Science Scale | Almost Never | About Monthly | .033 | .046 | 1.00 |
| | | About Weekly | .184 | .072 | .032 |
| | About Monthly | About Weekly | .151 | .077 | .154 |
| Student Confident in Math Scale | Almost Never | About Monthly | .235 | .051 | .000 |
| | | About Weekly | .357 | .079 | .000 |
| | About Monthly | About Weekly | .121 | .085 | .455 |
| Student Confident in Science Scale | Almost Never | About Monthly | .173 | .048 | .001 |
| | | About Weekly | .326 | .074 | .000 |
| | About Monthly | About Weekly | .153 | .080 | .169 |

Again, for the Math Assessment covariates, when using the Gender independent variable, only Plausible Values #3, and #5 were used, and when using the Race independent variable, only Plausible Value #1 was used. Similarly, when using the Gender independent variable and the Science Assessment covariate, only Plausible Value #4 was used, and when using the Race independent variable, only Plausible Value #1 was used.

Table 27

*ANOVA Investigation into the Independence between the Dependent Variable (Students Bullied at School Scale) and Math and Science Assessment Scores*

| Covariate Plausible Value | | Sum of Squares | Df | Mean Square | F | Sig |
|---|---|---|---|---|---|---|
| Math Plausible Value #1 | Between Groups | 62.37 | 1 | 62.37 | 16.88 | .000 |
| | Within Groups | 38226.49 | 10350 | 3.70 | | |
| | Total | 38288.85 | 10350 | | | |
| Math Plausible Value #3 | Between Groups | 32.65 | 1 | 32.65 | 8.83 | .003 |
| | Within Groups | 38256.20 | 10350 | 3.70 | | |
| | Total | 38288.85 | 10350 | | | |
| Math Plausible Value #5 | Between Groups | 36.97 | 1 | 36.97 | 10.00 | .002 |
| | Within Groups | 38251.88 | 10350 | 3.70 | | |
| | Total | 38288.85 | 10350 | | | |
| Science Plausible Value #1 | Between Groups | 5.50 | 1 | 5.50 | 1.49 | .223 |
| | Within Groups | 38283.35 | 10350 | 3.70 | | |
| | Total | 38288.85 | 10350 | | | |
| Science Plausible Value #4 | Between Groups | 7.32 | 1 | 7.32 | 1.98 | .160 |
| | Within Groups | 38281.53 | 10350 | 3.70 | | |
| | Total | 38288.85 | 10350 | | | |

Table 27 presents the ANOVA results between the Students Bullied at School Scale and those covariates. Table 28 reports the descriptive characteristics of the Student Bullied at School Index, while Table 29 illustrates the ANOVA results comparing students' math and science achievement and their responses on the Student Bullied at School Index. Finally, Table 30 reports the Bonferroni-adjusted post-hoc pairwise comparisons generated following the ANOVA calculations.

When the Student Bullied at School Scale is compared to the math and science assessment score covariates, results indicate that students' performance on the math assessment is significantly related to their reports of their frequency of bullying victimization (Math Plausible Value #1 = $F(1, 10350) = 16.88$; $p < .001$; Math Plausible Value #3 = $F(1, 10350) = 8.83$; $p = .003$; Math Plausible Value #5 = $F(1, 10350) =$

Table 28

*Demographic Data for Students Bullied at School Index by Math and Science Achievement Scores*

| Demographic (N) | Math Plausible Value #1 | | Math Plausible Value #3 | | Math Plausible Value #5 | | Science Plausible Value #1 | | Science Plausible Value #4 | |
|---|---|---|---|---|---|---|---|---|---|---|
| | *M* | *SD* | *M* | *SD* | *M* | *SD* | *M* | *SD* | *M* | *SD* |
| Almost Never (6,530) | 512.4 | 75.4 | 513.3 | 76.6 | 512.6 | 76.0 | 525.4 | 80.3 | 525.0 | 79.8 |
| About Monthly (2,880) | 509.2 | 74.0 | 511.1 | 74.1 | 510.6 | 75.3 | 527.4 | 80.4 | 527.1 | 79.3 |
| About Weekly (940) | 494.2 | 78.3 | 497.4 | 79.0 | 495.8 | 78.1 | 517.8 | 85.7 | 517.3 | 84.9 |
| Total | 509.9 | 75.4 | 511.3 | 76.2 | 510.5 | 76.1 | 525.3 | 80.9 | 524.9 | 80.2 |

10.00; $p = .002$). However, students' reports of bullying victimization were not found to be significantly related to their performance on the science assessment (Science Plausible Value #1 = $F(1, 10350) = 16.88$; $p < .001$; Science Plausible Value #4 = $F(1, 10350) = 1.98$; $p = .160$).

When the Index version of the Students Bullied at School variable was used as the dependent variable, however, students whose reports fell in different nominal groups of bullying victimization frequency demonstrated significantly different scores on both the math (Math Plausible Value #1 = $F(2, 10350) = 24.24$; $p < .001$; Math Plausible Value #3 = $F(2, 10350) = 17.93$; $p < .001$; Math Plausible Value #5 = $F(2, 10350) = 20.02$; $p < .001$) and science (Science Plausible Value #1 = $F(2, 10350) = 4.91$; $p = .007$; Science Plausible Value #4 = $F(2, 10350) = 5.32$; $p = .005$) assessments. Bonferroni-adjusted

Table 29

*ANOVA Investigation into Differences in the Student Bullied at School Index and Math and Science Achievement Scores*

| Covariate Plausible Value | | Sum of Squares | Df | Mean Square | F | Sig |
|---|---|---|---|---|---|---|
| Math Plausible Value #1 | Between Groups | 274737.8 | 2 | 137368.9 | 24.24 | .000 |
| | Within Groups | 58616970.3 | 10350 | 5666.2 | | |
| | Total | 58891708.1 | 10350 | | | |
| Math Plausible Value #3 | Between Groups | 207704.2 | 2 | 103852 | 17.93 | .000 |
| | Within Groups | 59935265.3 | 10350 | 5793.65 | | |
| | Total | 60142969.6 | 10350 | | | |
| Math Plausible Value #5 | Between Groups | 231007.1 | 2 | 115503.5 | 20.02 | .000 |
| | Within Groups | 59694874.6 | 10350 | 5770.4 | | |
| | Total | 59925881.7 | 10350 | | | |
| Science Plausible Value #1 | Between Groups | 64250.4 | 2 | 32125.2 | 4.91 | .007 |
| | Within Groups | 67627272.2 | 10350 | 6537.2 | | |
| | Total | 67691522.6 | 10350 | | | |
| Science Plausible Value #4 | Between Groups | 68329.5 | 2 | 34164 | 5.32 | .005 |
| | Within Groups | 66431531.0 | 10350 | 6421.6 | | |
| | Total | 66499860.5 | 10350 | | | |

post-hoc pairwise comparisons show that students who reported a bullying frequency rate of “About Weekly” earned significantly lower scores on the math assessment than students who reported “About Monthly” or “Almost Never” rates of bullying victimization. The mean differences between the scores suggest that students who reported “Almost Never” experiencing bullying victimization earned the highest scores on the math assessment, with students who reported “About Monthly” rates of victimization earning slightly lower scores; however, these scores were not significantly different. On the science assessment, students who reported the highest frequency of bullying victimization and fell in the “About Weekly” group also earned significantly lower scores than students who reported being bullied less frequently. The mean

Table 30

*Mean Difference Results of Bonferroni-Adjusted Pairwise Post-Hoc Analyses between Scores on Student Bullied at School Index and Math and Science Assessment Score Covariates*

| Plausible Value Covariate | (I) Student Bullied at School Index Rating | (J) Student Bullied at School Index Rating | Mean Difference (I-J) | Std. Error | Sig. |
|---|---|---|---|---|---|
| Math Plausible Value #1 | Almost Never | About Monthly | 3.25 | 1.68 | .161 |
| | | About Weekly | 18.25 | 2.63 | .000 |
| | About Monthly | About Weekly | 15.00 | 2.83 | .000 |
| Math Plausible Value #3 | Almost Never | About Monthly | 2.28 | 1.70 | .543 |
| | | About Weekly | 15.91 | 2.66 | .000 |
| | About Monthly | About Weekly | 13.64 | 2.86 | .000 |
| Math Plausible Value #5 | Almost Never | About Monthly | 2.04 | 1.70 | .692 |
| | | About Weekly | 16.79 | 2.65 | .000 |
| | About Monthly | About Weekly | 14.75 | 2.86 | .000 |
| Science Plausible Value #1 | Almost Never | About Monthly | -2.00 | 1.81 | .804 |
| | | About Weekly | 7.51 | 2.82 | .024 |
| | About Monthly | About Weekly | 9.51 | 3.04 | .005 |
| Science Plausible Value #4 | Almost Never | About Monthly | -2.18 | 1.79 | .672 |
| | | About Weekly | 7.64 | 2.80 | .019 |
| | About Monthly | About Weekly | 9.82 | 3.01 | .003 |

difference results of the science assessment suggest, however, that students who reported "About Monthly" rates of bullying victimization earned the highest scores, followed by students who reported bullying victimization "About Never;" although these differences were not statistically significant.

## Main Analyses

### Research Question 1

Does a student's identification with a specific race impact his or her experience of peer bullying victimization after controlling for the student's ability, enjoyment, and confidence in mathematics?

Hypothesis 1: After controlling for his or her mathematics ability, confidence, and enjoyment, a racial minority student will report higher frequencies of peer bullying victimization than a White student.

$H_0$ = there is no difference between the bullying rates of racial minority students and White students ($\mu_0 = \mu_1$)

$H_1$ = racial minority students experience significantly higher rates of bullying victimization than White students ($\mu_0 < \mu_1$; $\alpha < .05$)

As described, an ANCOVA was completed in order to answer this research question. The results of the ANCOVA are presented in Table 31. Results of this analysis indicate that a student's race is significantly related to his or her reports of the frequency of bullying victimization, $F(6, 10340) = 8.68$; $p < .001$, after controlling for the student's math ability score, his or her reports of enjoyment in math, and his or her confidence in math. While this is a significant result, the model tested, which included both the independent variable and the three covariates, is shown to have a very small effect size

Table 31

*ANCOVA Results Analyzing Students' Race, Math-Related Covariates, and Reports of Bulling Victimization Frequency*

| Source | Type III Sum of Squares | Df | Mean Square | F | Sig | Partial Eta Squared |
|---|---|---|---|---|---|---|
| Corrected Model | 338.07 | 9 | 37.56 | 10.23 | .000 | .009 |
| Intercept | 13429.59 | 1 | 13429.59 | 3658.29 | .000 | .261 |
| Math Assessment Plausible Value #1 | 35.52 | 1 | 35.52 | 9.68 | .002 | .001 |
| Student Likes Learning Math | .04 | 1 | .04 | .01 | .917 | .000 |
| Student Confident in Math | 48.40 | 1 | 48.40 | 13.18 | .000 | .001 |
| Race | 191.16 | 6 | 31.86 | 8.68 | .000 | .005 |
| Error | 37950.79 | 10340 | 3.67 | | | |
| Corrected Total | 38288.85 | 10350 | | | | |

($\eta^2 = .009$). This effect size demonstrates that the model explains only 0.9% of the variance in students' ratings on the Student Bullied at School Scale. The result of the ANCOVA also represents an increase in the *F* statistic and in the effect size determined by the ANOVA analysis comparing the mean reports of each racial demographic group on the Student Bullied at School Scale that is presented in Table 9, $F(6, 10340) = 6.84$; $p < .001$; $\eta^2 = .004$. However, the increase in the new model's ability to explain the variance in the dependent variable with the addition of the covariates represents a trivial amount.

A closer examination of the covariates used in the updated model shows that the Math Assessment Plausible Value, $F(1, 10340) = 9.68$; $p = .002$) and Student Confident in Math, $F(6, 10340) = 13.18$; $p < .001$) are significantly related to the dependent variable

Table 32

*Parameter Estimates Calculated During ANCOVA Procedure Using Race and Math-Related Covariates, Including Plausible Value #1*

| Parameter | *B* | Std. Error | *t* | Sig. | Partial Eta Squared |
|---|---|---|---|---|---|
| Intercept | 8.98 | .17 | 52.91 | .000 | .213 |
| Math Assessment Plausible Value #1 | .00 | .00 | 3.11 | .002 | .001 |
| Student Likes Learning Math | -.00 | .01 | -.10 | .917 | .000 |
| Student Confident in Math | .04 | .01 | 3.63 | .000 | .001 |
| White | .13 | .09 | 1.56 | .119 | .000 |
| Black | .25 | .10 | 2.54 | .011 | .001 |
| Hispanic | .44 | .09 | 4.89 | .000 | .002 |
| Asian | .27 | .12 | 2.21 | .027 | .000 |
| Native American | .02 | .20 | .10 | .924 | .000 |
| Pacific Islander | .33 | .21 | 1.59 | .112 | .000 |
| Two or More Races | 0 | | | | |

of the Student Bullied at School Scale. The Student Likes Learning Math covariate was not found to be significantly related to the outcome variable, $F(6, 10340) = .40$; $p = .917$). The beta values, which indicate the magnitude and direction of the relationship between each covariate and each racial demographic group, are presented in Table 32. These values indicate that the Math Assessment Plausible Value #1 ($B < .01$; $t(10340) = 3.11$; $p = .002$) and Student Confident in Math Scale covariates ($B = .04$; $t(10340) = 3.63$; $p < .001$) both have a significant, positive relationship with the outcome variable, meaning that as students' math ability and math confidence increase, so, too, does their reports of the frequency of bullying victimization.

Finally, Table 33 presents the Bonferroni-adjusted pairwise comparisons between the adjusted means of each racial demographic group's ratings on the Student Bullied at

Table 33

*Mean Difference Results Using Updated Group Means of Bonferroni-Adjusted Pairwise Post-Hoc Analyses between Students' Race and Responses on the Student Bullied at School Scale*

| Scale | (I) Race | (J) Race | Mean Difference (I-J) | Std. Error | Sig. |
|---|---|---|---|---|---|
| Student Bullied at School Scale | White | Black | -.188 | .063 | 1.00 |
| | | Hispanic | -.307 | .047 | .000 |
| | | Asian | -.132 | .093 | 1.00 |
| | | Native American | .116 | .179 | 1.00 |
| | | Pacific Islander | -.200 | .196 | 1.00 |
| | | 2 or more races | .144 | .086 | 1.00 |
| | Black | Hispanic | -.189 | .063 | .088 |
| | | Asian | -.015 | .016 | 1.00 |
| | | Native American | .233 | .185 | 1.00 |
| | | Pacific Islander | -.082 | .201 | 1.00 |
| | | 2 or more races | .252 | .099 | .231 |
| | Hispanic | Asian | .174 | .098 | 1.00 |
| | | Native American | .422 | .180 | .404 |
| | | Pacific Islander | .107 | .197 | 1.00 |
| | | 2 or more races | .441 | .090 | .000 |
| | Asian | Native American | .248 | .199 | 1.00 |
| | | Pacific Islander | -.067 | .214 | 1.00 |
| | | 2 or more races | .267 | .121 | .571 |
| | Native American | Pacific Islander | -.315 | .262 | 1.00 |
| | | 2 or more races | .019 | .195 | 1.00 |
| | Pacific Islander | 2 or more races | .334 | .210 | 1.00 |

School Scale after controlling for the math-related covariates in the model. When comparing these adjusted mean differences to the original mean differences displayed in Table 10, the results are remarkably unchanged. With this updated data, the same patterns of results persist, with Hispanic students reporting significantly higher rates of bullying victimization than White students and Multi-racial students. No other significant differences were found between the reports of the different racial groups.

**Research Question 2**

Does a student's identification with a specific race impact his or her experience of peer bullying victimization after controlling for the student's ability, enjoyment, and confidence in science?

Hypothesis 1: After controlling for his or her science ability, confidence, and enjoyment, a racial minority student will report higher frequencies of peer bullying victimization than a White student.

$H_0$ = there is no difference between the bullying rates of racial minority students and White students ($\mu_0 = \mu_1$)

$H_1$ = racial minority students experience significantly higher rates of bullying victimization than White students ($\mu_0 < \mu_1$; $\alpha < .05$)

Another ANCOVA procedure was conducted in order to answer this research question. Results of this analysis are presented in Table 34. Results show that students' race was, again, significantly related to reports of the frequency of bullying victimization, $F(6, 10340) = 7.05$; $p < .001$) after controlling for students' science ability score, their reports of enjoyment in science, and their confidence in science. However, this significant result is also found to have a very small effect size ($\eta^2 = .008$), and, therefore, explains only 0.8% of the variance in the Student Bullied at School Scale. This model also represents an increased $F$ statistic and effect size when compared to the previous model's results, $F(6, 10340) = 6.84$; $p < .001$; $\eta^2 = .004$), but this increase is even smaller

Table 34

*ANCOVA Results Analyzing Students' Race, Science-Related Covariates, and Reports of Bulling Victimization Frequency*

| Source | Type III Sum of Squares | Df | Mean Square | F | Sig | Partial Eta Squared |
|---|---|---|---|---|---|---|
| Corrected Model | 308.70 | 9 | 34.30 | 9.34 | .000 | .008 |
| Intercept | 15837.30 | 1 | 15837.30 | 4310.83 | .000 | .294 |
| Science Assessment Plausible Value #1 | 9.60 | 1 | 9.60 | 2.61 | .106 | .000 |
| Student Likes Learning Science | 16.49 | 1 | 16.49 | 4.49 | .034 | .000 |
| Student Confident in Science | 137.01 | 1 | 137.01 | 37.29 | .000 | .004 |
| Race | 155.49 | 6 | 25.91 | 7.05 | .000 | .004 |
| Error | 37980.15 | 10340 | 3.67 | | | |
| Corrected Total | 38288.85 | 10350 | | | | |

than the improvement demonstrated in the previous ANCOVA that used math-related covariates.

When the science-related covariates are examined, the Student Likes Learning Science, $F(1, 10340) = 4.49$; $p = .034$, and Student Confident in Science, $F(1, 10340) = 37.29$; $p < .001$, were found to be significantly related to the dependent variable scale. The Science Assessment Plausible Value #1 covariate, however, was not found to be significantly related to the outcome variable, $F(1, 10340) = 2.61$; $p = .106$). These results are presented in Table 35.

The beta values presented in Table 35 indicate that the Student Likes Learning Science Scale has a significant negative relationship with the Student Bullied at School Scale ($B = -.03$; $t(10340) = -2.12$; $p = .034$), which shows that as students' enjoyment in science decreases, their reports of bullying victimization increases. The Student

Table 35

*Parameter Estimates Calculated During ANCOVA Procedure Using Race and Science-Related Covariates, Including Plausible Value #1*

| Parameter | *B* | Std. Error | *t* | Sig. | Partial Eta Squared |
|---|---|---|---|---|---|
| Intercept | 9.59 | .17 | 56.65 | .000 | .237 |
| Science Assessment Plausible Value #1 | .00 | .00 | -1.62 | .106 | .000 |
| Student Likes Learning Science | -.03 | .01 | -2.12 | .034 | .000 |
| Student Confident in Science | .08 | .01 | 6.11 | .000 | .004 |
| White | .15 | .09 | 1.76 | .078 | .000 |
| Black | .21 | .10 | 2.06 | .038 | .000 |
| Hispanic | .41 | .09 | 4.56 | .000 | .002 |
| Asian | .34 | .12 | 2.85 | .004 | .001 |
| Native American | -.03 | .20 | -.17 | .865 | .000 |
| Pacific Islander | .28 | .21 | 1.34 | .180 | .000 |
| Two or More Races | 0 | | | | |

Confidence in Science Scale, however, has a significant positive relationship with the bullying outcome variable ($B = .08$; $t(10340) = 6.11$; $p < .001$).

Similar to the previous analysis, Table 36 presents the Bonferroni-adjusted pairwise comparisons between the adjusted means of each racial demographic group's ratings on the Student Bullied at School Scale after controlling for the science-related covariates in the model. When comparing these adjusted mean differences to the original mean differences displayed in Table 10, the results continue to be generally unchanged.

After controlling for the science-related covariates, Hispanic students continue to report significantly higher rates of bullying victimization than both White and Multi-racial students. However, with this updated data, Hispanic students were also found to

Table 36

*Mean Difference Results Using Updated Group Means of Bonferroni-Adjusted Pairwise Post-Hoc Analyses between Students' Race and Responses on the Student Bullied at School Scale*

| Scale | (I) Race | (J) Race | Mean Difference (I-J) | Std. Error | Sig. |
|---|---|---|---|---|---|
| Student Bullied at School Scale | White | Black | -.054 | .064 | 1.00 |
| | | Hispanic | -.261 | .048 | .000 |
| | | Asian | -.191 | .092 | .807 |
| | | Native American | .185 | .179 | 1.00 |
| | | Pacific Islander | -.131 | .196 | 1.00 |
| | | 2 or more races | .152 | .086 | 1.00 |
| | Black | Hispanic | -.206 | .066 | .036 |
| | | Asian | -.136 | .105 | 1.00 |
| | | Native American | .239 | .185 | 1.00 |
| | | Pacific Islander | -.077 | .201 | 1.00 |
| | | 2 or more races | .206 | .099 | .801 |
| | Hispanic | Asian | .070 | .097 | 1.00 |
| | | Native American | .445 | .180 | .284 |
| | | Pacific Islander | .130 | .197 | 1.00 |
| | | 2 or more races | .412 | .090 | .000 |
| | Asian | Native American | .375 | .198 | 1.00 |
| | | Pacific Islander | .060 | .214 | 1.00 |
| | | 2 or more races | .342 | .120 | .093 |
| | Native American | Pacific Islander | -.316 | .262 | 1.00 |
| | | 2 or more races | -.033 | .195 | 1.00 |
| | Pacific Islander | 2 or more races | .282 | .211 | 1.00 |

report significantly higher rates of bullying victimization compared to Black students as well.

**Research Question 3**

Does a student's gender impact his or her experience of peer bullying victimization after controlling for the student's ability, enjoyment, and confidence in mathematics?

Hypothesis 1: After controlling for her mathematics ability, confidence, and enjoyment, a female student will report higher frequencies of peer bullying victimization than a male student.

$H_0$ = there is no difference between the bullying rates of female students and male students ($\mu_0 = \mu_1$)

$H_1$ = female students experience significantly higher rates of bullying victimization than male students ($\mu_0 < \mu_1$; $\alpha < .05$)

Another ANCOVA was completed in order to answer this research question. The Student Bullied at School Scale continued to be the dependent variable, while the Gender variable was used as the independent variable. For this analysis, the Student Likes Learning Math and Student Confident in Math were both used as covariates, while two separate analyses were completed using Plausible Values #3 and #5 separately as additional covariates.

Results of the ANCOVA analyses indicate that, when using either Math Assessment Plausible Value #3, $F(1, 10340) = 7.49$; $p = .006$, or Plausible Value #5, $F(1, 10340) = 7.52$; $p = .006$, a student's gender is significantly related to the reported frequency of bullying victimization after controlling for math assessment score, students' enjoyment of math, and students' confidence in math. Table 37 further describes the results of this ANCOVA. While this is a statistically significant result, this model was found to have a very small effect size ($\eta^2 = .004$) and, therefore, only accounts for 0.4%

Table 37

*ANCOVA Results Analyzing Students' Gender, Math-Related Covariates and Reports of Bulling Victimization Frequency*

| Source | Type III Sum of Squares | Df | Mean Square | F | Sig | Partial Eta Squared |
|---|---|---|---|---|---|---|
| | Math Assessment Plausible Value #3 | | | | | |
| Corrected Model | 164.14 | 4 | 41.03 | 11.13 | .000 | .004 |
| Intercept | 16412.08 | 1 | 16412.08 | 4452.50 | .000 | .301 |
| Math Assessment Plausible Value #3 | .60 | 1 | .60 | .16 | .686 | .000 |
| Student Likes Learning Math | .56 | 1 | .56 | .15 | .697 | .000 |
| Student Confident in Math | 63.73 | 1 | 63.73 | 17.29 | .000 | .002 |
| Gender | 27.62 | 1 | 27.62 | 7.49 | .006 | .001 |
| Error | 38124.72 | 10340 | 3.69 | | | |
| Corrected Total | 38288.85 | 10350 | | | | |
| | Math Assessment Plausible Value #5 | | | | | |
| Corrected Model | 165.28 | 4 | 41.32 | 11.21 | .000 | .004 |
| Intercept | 16265.40 | 1 | 16265.40 | 4412.84 | .000 | .299 |
| Plausible Value #5 | 1.75 | 1 | 1.75 | .48 | .491 | .000 |
| Student Likes Learning Math | .62 | 1 | .62 | .17 | .683 | .000 |
| Student Confident in Math | 60.83 | 1 | 60.83 | 16.50 | .000 | .002 |
| Gender | 27.73 | 1 | 27.73 | 7.52 | .006 | .001 |
| Error | 38123.57 | 10340 | 3.69 | | | |
| Corrected Total | 38288.85 | 10350 | | | | |

of the variance in the Student Bullied at School Scale. This result does represent an increase in the *F* statistic and in the effect size found from the ANOVA between the Gender independent variable and the Student Bullied at School Scale dependent variable as reported in Table 9, $F(1, 10350) = 4.91$; $p = .027$; $\eta^2 = .0005$. However, this change is trivial given the statistical power of the analyses due to the large sample size.

Table 38

*Parameter Estimates Calculated During ANCOVA Procedure Using Gender and Math-Related Covariates, Including both Math Assessment Plausible Values #3 and #5*

| Parameter | *B* | Std. Error | *t* | Sig. | Partial Eta Squared |
|---|---|---|---|---|---|
| | Math Assessment Plausible Value #3 | | | | |
| Intercept | 9.45 | .14 | 65.51 | .000 | .293 |
| Math Assessment Plausible Value #3 | .00 | .00 | .41 | .686 | .000 |
| Student Likes Learning Math | .01 | .01 | .39 | .697 | .000 |
| Student Confident in Math | .05 | .01 | 4.16 | .000 | .002 |
| Female | .10 | .04 | 2.74 | .006 | .001 |
| Male | 0 | | | | |
| | Math Assessment Plausible Value #5 | | | | |
| Intercept | 9.42 | .15 | 65.11 | .000 | .291 |
| Math Assessment Plausible Value #5 | .00 | .00 | .69 | .491 | .000 |
| Student Likes Learning Math | .01 | .01 | .41 | .683 | .000 |
| Student Confident in Math | .05 | .01 | 4.10 | .000 | .002 |
| Female | .10 | .04 | 2.74 | .006 | .001 |
| Male | 0 | | | | |

Parameter estimates when using both Plausible Value #3 and #5 are displayed in Table 38. These estimates indicate that the Student Confident in Math scale was the only covariate that was significantly related to the Student Bullied at School dependent variable (Plausible Value #3 = $t(10340) = 4.16$; $p < .001$; Plausible Value #5 = $t(10340) =$ $4.06$; $p < .001$). However, the beta values generated during both calculations are small and remain somewhat trivial (Plausible Value #3 = $B = .048$; Plausible Value #5 = $B =$ $.047$). Finally, as hypothesized, results show that female students report more frequent rates of bullying victimization on the Students Bullied at School Scale than did male students, $t(10340) = 2.74$; $p = .006$.

The similarity of the results generated between the analyses when using the Math Assessment Plausible Values #3 and #5 do provide further support to the belief that the ANCOVA procedure is robust to the violation of assumptions described above. During the discussion regarding the independence of the covariates and the independent variable, Plausible Value #3 was shown to not violate the assumption, while Plausible Value #5 did violate the assumption. Despite this difference, the outcomes of the ANCOVA analysis using the two Plausible Values have resulted in similar conclusions.

**Research Question 4**

Does a student's gender impact his or her experience of peer bullying victimization after controlling for the student's ability, enjoyment, and confidence in science?

Hypothesis 1: After controlling for her science ability, confidence, and enjoyment, a female student will report higher frequencies of peer bullying victimization than a male student.

$H_0$ = there is no difference between the bullying rates of female students and male students ($\mu_0 = \mu_1$)

$H_1$ = female students experience significantly higher rates of bullying victimization than male students ($\mu_0 < \mu_1$; $\alpha < .05$)

The final ANCOVA calculation was completed in order to answer this research question. The Student Bullied at School Scale continued to be the dependent variable, while the Gender variable was, again, used as the independent variable. For this analysis, the Student Likes Learning Science and Student Confident in Science, and Plausible

Table 39

*ANCOVA Results Analyzing Students' Gender, Science-Related Covariates, and Reports of Bulling Victimization Frequency*

| Source | Type III Sum of Squares | Df | Mean Square | F | Sig | Partial Eta Squared |
|---|---|---|---|---|---|---|
| Corrected Model | 182.85 | 4 | 45.71 | 12.41 | .000 | .005 |
| Intercept | 18786.73 | 1 | 18786.73 | 5099.23 | .000 | .330 |
| Plausible Value #4 | 39.86 | 1 | 39.86 | 10.82 | .001 | .001 |
| Student Likes Learning Science | 10.07 | 1 | 10.07 | 2.73 | .098 | .000 |
| Student Confident in Science | 130.45 | 1 | 130.45 | 35.41 | .000 | .003 |
| Gender | 24.59 | 1 | 24.59 | 6.67 | .010 | .001 |
| Error | 38106.00 | 10340 | 3.68 | | | |
| Corrected Total | 38288.85 | 10350 | | | | |

Value #4 from the Science Assessment were all used as covariates.

ANCOVA results indicate that, after controlling for students' enjoyment, confidence, and level of achievement in science, students' gender is significantly related to their reports of the frequency of bullying victimization, $F(1, 10340) = 6.67$; $p = .010$ (Table 39). While this is a statistically significant finding, this model accounts for a very small amount of the variance in reports of bullying victimization frequency ($\eta^2 = .005$). The ANCOVA results, also, demonstrate an increase in statistical significance and the amount of variance explained over the ANOVA procedure that was discussed above in Table 9, $F(1, 10350) = 4.91$; $p = .027$; $\eta^2 = .0005$. Again, the difference in these results is somewhat trivial.

Results do, however, continue to support the hypothesis that female students reported significantly higher rates of bullying victimization than did male students, $t(10340) = 2.58$; $p = .010$, after controlling for the science-related covariates. Further

Table 40

*Parameter Estimates Calculated During ANCOVA Procedure Using Gender and Science-Related Covariates, Including Plausible Value #4*

| Parameter | *B* | Std. Error | *t* | Sig. | Partial Eta Squared |
|---|---|---|---|---|---|
| | Science Assessment Plausible Value #4 | | | | |
| Intercept | 9.93 | .14 | 69.47 | .000 | .318 |
| Plausible Value #4 | -.00 | .00 | -3.29 | .001 | .001 |
| Student Likes Learning Science | -.02 | .01 | -1.65 | .098 | .000 |
| Student Confident in Science | .07 | .01 | 5.95 | .000 | .003 |
| Female | .10 | .04 | 2.58 | .010 | .001 |
| Male | 0 | | | | |

descriptions of the parameter estimates are presented in Table 40. Finally, further examination of the covariates used in the ANCOVA calculates indicate that both the Science Assessment Plausible Value #4, $t(10340) = -3.29$; $p = .001$, and the Student Confident in Science, $t(10340) = 5.95$; $p < .001$, covariates were significantly related to students' reports on the Student Bullied at School Scale. Beta values indicate that, while responses on the Student Confident in Science scale were positively related to students' reports of bullying victimization frequency, $B = .07$), students' scores on the Science Assessment were negatively related to reports of bullying victimization frequency, $B < -.01$). However, this negative relationship was extremely small.

**Summary**

After preparing the original TIMSS 2011 dataset for use, a final sample of 10,350 students was available to answer the established research questions. This final sample consisted of 51% female students, 50% White students, 26% Hispanic students, and 12%

Black students, with other racial minority groups comprising 5% or less of the total sample.

First, the assumptions of the Analysis of Covariance (ANCOVA) procedure were considered in order to determine if the findings of the subsequent analyses could be clearly interpreted. After checking the assumptions, it was concluded that the data from the TIMSS 2011 dataset did violate the independence of covariates and independent variables assumption, as well as the homogeneity of regression slopes assumption. Specifically, these calculations found that students of different gender and racial groups demonstrated significantly different patterns of responses on the covariate scales. Female students were shown to report significantly higher rates of bullying victimization, and significantly lower levels of enjoyment, confidence, and performance scores related to both math and science than did their male peers. While there was some variability in the race-based results, in general, Hispanic students reported significantly higher rates of bullying victimization than other racial groups, while Asian and White students reported the highest levels of enjoyment, confidence, and performance in math and science. Because the data violate the assumptions of an ANCOVA, the results obtained should be interpreted with some caution.

Next, a series of preliminary analyses were conducted in order to better understand the relationships between students' reports of bullying victimization and different groups' levels of performance in math and science. These analyses found that female and male students appear to have qualitatively different relationships between math and science performance and bullying victimization. When math performance was examined, female students were found to report a statistically significant positive correlation between math assessment scores and rate of bullying victimization, while the relationship between bullying and math performance was non-significant for male students. When science performance was examined, female students continued to report a positive relationship between bullying and science assessment score; however, this correlation did not reach statistical significance. Male students, however, reported a statistically significant negative correlation between bullying victimization and science assessment score. These same relationships were also calculated for each racial demographic group, and White and Hispanic students were found to report statistically significant positive correlations between bullying victimization and math assessment score. However, these were the only statistically significant findings in terms of racial groups.

Another set of preliminary analyses was completed. These calculations were related to the academic performance and perceptions of students who reported different frequencies of bullying victimization. For these analyses, the Student Bullied at School Index was examined rather than the Student Bullied at School Scale, which was used for all other analyses. The Student Bullied at School Index created three distinct groups of students who were rated as experiencing peer bullying victimization either “About Weekly,” “About Monthly,” or “Almost Never.” After comparing these three groups’ levels of academic performance and academic perceptions, those students who reported experiencing bullying victimization “About Weekly” showed significantly lower levels of math and science performance, and reported significantly lower rates of enjoyment and confidence in both math and science than students who were bullied less frequently.

Finally, each of the four main analyses were completed using the ANCOVA procedure. First, when controlling for the math performance, enjoyment, and confidence variables, there were found to be statistically significant racial differences in terms of reports of bullying victimization. Post-hoc pairwise comparisons concluded that Hispanic students' reports of bullying victimization were significantly higher than the reports made by White and Multi-Racial students. Second, when controlling for the science performance, enjoyment, and confidence variables, similar results were gathered, with Hispanic students' demonstrating significantly higher rates of bullying victimization than White, Multi-Racial, and Black students. Third, when, again, controlling for the math performance, enjoyment, and confidence variables, female students were shown to report significantly higher rates of bullying victimization than their male peers. Finally, this same pattern of results, with female students reporting higher rates of bullying than male students, arose after controlling for the science performance, enjoyment, and confidence variables.

While each of the results obtained by the main analyses did reach a level of statistical significance, each of the developed models reported a very small effect size ($\eta^2 < .01$). This small effect size indicates that the differences found between the racial and gender groups bears little practical significance in students' experience of bullying victimization. Furthermore, the addition of the covariates to the models did little to increase the models' ability to explain the variance in peer bullying victimization above and beyond the level of explanation provided by the independent variables of race and gender. Therefore, while it was appropriate to reject the null hypotheses established for the four main analyses, these results do not provide conclusive evidence regarding the relationships between gender, race, bullying victimization, and academic performance.

# CHAPTER 5 – Discussion

## Summary of Findings

### Description of Study

In the current investigation, the researcher sought to determine if female and/or racial minority students experienced more frequent rates of bullying victimization than their male or White peers after controlling for students' demonstrated levels of math or science skills, reported level of enjoyment in math or science, and reported level of confidence in math or science. It was originally hypothesized that female and racial minority students would experience more frequent bullying victimization. The reasoning behind this hypothesis was largely related to previous literature that had described how students who were different in some way from their peer group (Sweeting & West, 2001), including demonstrating differences in academic performance (Bishop et al., 2004), are at a greater risk for bullying victimization. Specifically, because of the consistent academic achievement gap that has been measured between female and male students, and between racial minority and White students (Ladson-Billings, 2006), those female or minority students who perform higher academically than their male or White peers, will appear noticeably different and, therefore, may be at an increased risk of bullying victimization (Peguero & Williams, 2011; Williams & Peguero, 2013).

In order to answer these research questions, the restricted use version of the TIMSS 2011 dataset was obtained from the National Center for Educational Statistics (NCES). The TIMSS 2011 dataset was selected due to its inclusion of Math and Science Assessments, which provided valid estimates of students' actual abilities in these academic areas. The dataset also included a robust Student Questionnaire, which asked students to report how they perceived both math and science coursework in terms of enjoyment and confidence, and also included questions related to the frequency at which they experience peer bullying victimization while at school. Another positive attribute of the TIMSS 2011 dataset was its rigorous sampling procedure and large sample size. The demographic characteristics of the TIMSS 2011 participants closely matched the overall demographics of the United States at large. The sample size, which consisted of over 10,000 students, also ensured that conclusions generated with the dataset would be ecologically valid.

After obtaining the dataset, it was determined that an Analysis of Covariance (ANCOVA) procedure would be an effective method of answering the research questions. When using the ANCOVA, students' responses on the Student Bullied at School Scale were used as the dependent variable given the study's purpose of better understanding the impact of academic functioning on the frequency of peer bullying victimization. Next, the student-reported variables of Gender and Race were selected as the study's independent variables due to the stated interest in understanding the differences in bullying victimization between gender- and race-based groups. Finally, in order to maximize the group differences in reported bullying victimization, students' performances on the Math and Science Assessments, reports on the Student Likes Learning Math and Science Scales, and reports on Student Confident Learning Math and Science Scales were selected as covariates.

Early in the statistical analysis process, however, it was determined that the variables selected for this procedure violated the assumptions of the ANCOVA procedure as described by Field (2013). Specially, the covariates were shown to not be independent of the two treatment variables of gender and race. This indicates that the covariates and the independent variable account for an unacceptably-large shared portion of the variance in the dependent variable. Therefore, the interpretation of the supposed impact of the independent variable on the dependent variable is somewhat uncertain as it is unclear how much variance in the dependent variable is uniquely explained by the Gender and Race variables.

Furthermore, the data was also shown to violate the assumption related to the homogeneity of the regression slopes of the covariates. This assumption states that the each of the covariates' regression slopes should be equal across each of the treatment groups (e.g., gender and race) to ensure that each covariate produces similar effects on each group. Because the data violates this assumption, the effects of the covariates do not appear to be consistent across each of the gender and racial groups. This further complicates the interpretation of the main ANCOVA analyses.

## Preliminary Analyses

**Student Questionnaire and academic assessment results.** Prior to completing the main analyses, several other calculations were completed in order to better understand the relationships between gender, race, academic performance, and peer bullying victimization. The first set of analyses explored the gender differences in both the bullying and academic variables. When an Analysis of Variance (ANOVA) procedure compared the mean reports of bullying victimization gathered from male and female students, results indicated that female students reported significantly higher rates of peer bullying victimization than did their male peers. While this finding was statistically significant, a student's gender accounted for less than 1% of the variance of in the bullying victimization variable, and, therefore, appears to have little noticeable impact in students' reports of being bullied.

Significant differences were also found between male and female students' reported levels of enjoyment and confidence in both math and science. Male students reported significantly more positive perceptions across each of these ratings, including liking of math, liking of science, confidence in math, and confidence in science. A closer examination of the *F* statistics that were generated during these ANOVA calculations indicates that the differences in mean reports between the genders is much larger on question items related to science than on items related to math. This indicates that the difference between male and female students' enjoyment and confidence in science is noticeably greater than the difference between male and female students' enjoyment and confidence in math.

Not only were male students found to report significantly more positive associations with science and math, but male students also demonstrated significantly higher levels of performance in both subjects as measured by the Math and Science Assessments. Again, this significant difference is most consistent and noticeably larger when examining results of the Science Assessment, as male students significantly outperformed female students on all five of the Plausible Values that were calculated. On the Math Assessment, while male students earned higher mean scores on each of the Plausible Values generated, only two of the five Plausible Value scores were found to demonstrate a mean difference that reached a level of statistical significance.

Preliminary racial differences were also examined. First, when an ANOVA was used to compare the means of each racial demographic groups' average reports of peer bullying victimization, results indicated that there was a statistically significant difference between the racial groups' reports. Follow-up analyses found that the overall significant result was due to the significantly discrepant mean differences between Hispanic students and White students, as well as between Hispanic students and Multi-Racial students. In both instances, Hispanic students reported significantly more frequent rates of bullying victimization. All other pairwise comparisons were not found to be significantly different. This significant result, however, continued to account for less than 1% of the variance in the Student Bullied at School Scale. This suggests that, similar to the findings related to students' gender, students' racial demographic status explains a very small amount of students' experiences of bullying victimization.

There continued to be significant differences between the racial demographic groups when students' enjoyment and confidence in both math and science were examined. Results gathered from the Student Likes Learning Math Scale indicate that Asian students reported the highest rate of enjoyment in math, and the average of this group's reports were significantly higher than the average reports of White, Hispanic, Native American, Pacific Islander, and Multi-Racial students. Black students reported the next highest level of enjoyment in math, and their reports were significantly higher than White, Hispanic, and Multi-racial students' reports. Finally, Hispanic students reported significantly higher rates of liking math than did White students.

On the Student Confident Learning Math Scale, Asian students reported the highest rates of confidence in math as well. Asian students' reports were found to be significantly higher than White, Black, Hispanic, Native American, Pacific Islander, and Multi-Racial students. Unlike the results on the liking math scale, White students reported the second highest rates of confidence in math; however, these reports were only significantly higher than Hispanic students' reports. Black students reported the third highest rates of math confidence, but again, this result was only significantly higher than the reports made by Hispanic students.

Results on the two scales related to students' perceptions related to science were more consistent than the results related to math. On the Student Likes Learning Science Scale, Asian students continued to report the most positive associations, and their reports were significantly higher than the reports made by White, Black, and Hispanic students. The reports of liking science made by White students were, then, found to be significantly higher than the reports made by Black and Hispanic students, as well. Asian students also reported the highest rates of confidence in science on the Student Confident Learning Science Scale, and those reports were significantly higher than the reports gathered from Black and Hispanic students. Both White and Multi-racial students also reported significantly higher rates of confidence with science than did Black and Hispanic students.

The mean results obtained on the Math and Science Assessments were also shown to vary significantly depending on students' reported racial demographic. On the Math Assessment, Asian students earned a significantly higher score than all other racial demographic groups on all five of the Plausible Values calculated. White students were then found to earn significantly higher scores than all other racial groups, except for Asian students, on all five Plausible Values. Multi-racial students then demonstrated the next highest level of math achievement, and earned significantly higher scores than Black, Hispanic, and Native American students on all five Plausible Values, and significantly higher scores than Pacific Islander students on four of the five Plausible Values. Finally, Hispanic students earned significantly higher math achievement scores than Black students on all five Plausible Values, and higher scores than Native American students on two of the five Plausible Values.

A similar pattern of results were found on the Science Assessment. Similar to the results of the Math Assessment, Asian and White students earned the highest scores, and the scores obtained by both of these groups were shown to be significantly higher than the scores obtained by all other racial demographic groups. However, unlike the results of the Math Assessment, the level of performance demonstrated by Asian and White students was remarkably similar. Multi-racial students earned the next highest average score, and that group's score was significantly higher than Black, Hispanic, Native American, and Pacific Islander students. Finally, Hispanic students were found to earn significantly higher scores than did Black students.

**Results related to bullying victimization.** In order to further explore the relationships between students' reports related to the frequency of bullying victimization, their academic performance, and their perceptions of enjoyment and confidence in math and science, additional ANOVA calculations were completed. During these analyses, the Student Bullied at School Index was used as the independent variable rather than the Student Bullied at School Scale, which was used in most analyses. This index variable allowed for the comparison of students who reported distinct levels of bullying victimization. These levels were labeled as either "Almost Never," "About Monthly," and "About Weekly." Each group's mean responses on the Student Questionnaire scales, as well as their scores on the Math and Science Assessments were then compared.

First, ANOVA results indicated that students who reported different frequencies of bullying victimization also reported significantly different levels of liking math and science and different levels of having confidence in math and science. Post-hoc comparisons between the three bullying frequency groups indicated that students who reported "Almost Never" experiencing bullying victimization also reported significantly higher levels of enjoyment in both science and math than did students who reported being bullied "About Weekly." When students' levels of confidence in math and science were compared, students who were bullied "Almost Never" reported significantly higher levels of confidence in both math and science than students who reported bullying victimization frequencies of both "About Weekly" and "About Monthly." These results indicate that students who experience more frequent peer bullying victimization also report lower levels of enjoyment and confidence in both math and science than students who are rarely bullied by their peers.

Next, a similar pattern of results was found when students' performance on the Math and Science Assessments were compared. Again, ANOVA results indicated that students who reported different rates of bullying victimization did earn significantly different scores on the academic skill measures. Post-hoc comparisons using each of the included Math and Science Plausible Values indicated that students who reported being bullied by their peers "About Weekly" earned significantly lower scores than their peers who reported being bullied "About Monthly" and "Almost Never." These results show that students who report being bullied more frequently demonstrate significantly lower levels of math and science skill than students who report being bullied less frequently.

In order to compare students' reports on the Student Bullied at School Scale and students' level of academic performance on the Math and Science Assessments, correlational and ANOVA calculations were completed. When completed using the total sample, there was found to be a significant positive relationship between students' reports of bullying victimization and their score on the Math Assessment. This suggests that, in general, as students' level of math ability increases, students tend to report more frequent rates of peer bullying victimization. However, there was a non-significant, and slightly negative relationship between students' reports of bullying victimization and their performance on the Science Assessment. This indicates that, in general, as students' science ability increases, they tend to demonstrate a slight decrease in the frequency of peer bullying victimization.

The correlations for each individual gender and racial demographic group were also calculated to further explore these relationships. When the correlations between bullying victimization and math ability were calculated for each gender, results indicated that female students reported a significant positive relationship between bullying victimization frequency and math ability, while the relationship between these two variables was non-significant, and almost nonexistent, for male students. This finding suggests that female students do tend to report more frequent experiences of peer bullying victimization as their math ability increases, but math ability has little to no impact on male students' reports of bullying victimization.

When each gender's bullying victimization reports and science abilities were compared, another interesting pattern emerged. While the correlation between these two variables was slightly negative when using the entire student sample, when each gender was considered separately, male students were shown to report a statistically significant, negative relationship between science ability and bullying victimization. Female students, however, report a non-significant, but positive, relationship between the two variables. This finding suggests that as male students' science ability increases, their reports of bullying victimization frequency tend to decrease. For female students, however, as their science ability increases, they continue to report a small tendency to be bullied more frequently.

Results of the correlational relationships between academic ability and bullying victimization are somewhat less clear when each racial demographic group is considered separately. When students' scores on the Math Assessment and their reports on the Student Bullied at School Scale were compared, only White and Hispanic students demonstrated a statistically significant, positive relationship between the two variables. Results obtained from Asian students led to a positive correlation that approached statistical significance, but did not achieve that level. The relationships between these two variables were also shown to be positive for Black, Pacific Islander, and Multi-racial students, but these relationships did not reach a level of statistical significance. Native American students, however, reported a negative relationship between math ability and bulling victimization frequency. These results suggest that, in general, students of each race tend to report higher rates of peer bullying victimization as their math ability increases. Native American students, however, report the opposite relationship, and tend to report less frequent experiences of bullying victimization as their math ability increases.

The relationship between students' reports of bullying victimization and science ability is less consistent across the different racial demographic groups. First, none of the correlations between the two variables for any of the racial groups reached a level of statistical significance. Furthermore, the direction of the relationship between the variables is inconsistent across racial groups. Three racial groups, White, Asian, and Pacific Islander, report a slight positive correlation, while four groups, Black, Hispanic, Native American, and Multi-racial, report small negative relationships. Given the lack of statistically significant results, and the lack of a priori hypotheses that would explain the variation in the direction of the relationships, these results should be interpreted with caution.

Overall, preliminary analyses indicate that there are several important whole sample-, gender-, and racial-based differences in students' tendencies to describe their frequency of bullying victimization, perceptions of math and science, and demonstrated skills in math and science. First, female students and Hispanic students were found to report significantly higher rates of bullying victimization than male students and students belonging to other racial demographic groups, respectively. Second, male students reported significantly higher levels of liking math and science and feeling confident in math and science than their female peers, while Asian students consistently reported the highest levels of enjoyment and confidence related to both math and science. Third, in terms of Math and Science Assessment performance, male and Asian students significantly outperformed female and other racial demographic groups, respectively, on the Math Assessment. Male students also outscored female students on the Science Assessment, while Asian students outperformed all other racial demographic groups, expect for White students, who earned roughly the same scores, on the Science Assessment.

Then, more significant relationships arose when students' academic perceptions and performances were compared to their reports of peer bullying victimization. When using the entire sample, students who reported experiencing more frequent episodes of bullying victimization reported significantly lower levels of enjoyment in both math and science, confidence in both math and science, and earned significantly lower scores on the Math and Science Assessments than did students who reported being bullied less frequently. Furthermore, correlational analyses indicated that, as their math and science scores increased, female students tended to report higher rates of bullying victimization. Male students, however, reported no relationship between math ability and frequency of bullying victimization, while there was a negative relationship between science ability and bullying victimization. When the sample was divided by racial demographic groups, most racial groups reported a positive correlation between math achievement score and frequency of bullying victimization, with White and Hispanic students' correlations reaching a level of statistical significance. The direction and strength of the relationships between science achievement score and bullying victimization frequency, however, was less consistent and interpretable.

## Main Analyses

**Research question 1.** The first research question asked if a student's identification with a specific race/ethnicity has an impact on his or her experience of peer bullying victimization after controlling for the student's ability, enjoyment, and confidence in mathematics. Based on previous literature, it was hypothesized that minority students would report higher frequencies of peer bullying victimization than White students. Results gathered by the completed ANCOVA indicated that a student's race was significantly related to his or her reports of bullying victimization. The ANCOVA model, which included the covariates related to students' math skills, math enjoyment, and math confidence, resulted in an increased difference between the racial groups, as well as an increased effect size when compared to the ANOVA analysis comparing the groups without including the covariates. This indicates that the math-related covariates did account for at least a small portion of the variance in students' reports of bullying victimization. Post hoc pairwise comparisons showed that this overall significant result was largely due to Hispanic students' significantly higher reports of peer bullying victimization than the reports made by White students and Multi-Racial students.

Based on these statistically significant results, it is concluded that the null hypothesis should be rejected for the first research question, as a minority student group was found to report significantly higher rates of bullying victimization. However, while the null hypothesis should be rejected, the gathered results do not appear to support the implicit hypothesis that all racial minority groups would experience more frequent rates of bullying victimization. Furthermore, the effect size of the model used during the ANCOVA was shown to account for a very small portion of the variance in the frequency of bullying. This suggests that, even though there is a statistically significant difference between the racial groups, there is little noticeable or practical difference between the groups.

**Research question 2.** The second research question asked if a student's identification with a specific race/ethnicity has an impact on his or her experience of peer bullying victimization after controlling for the student's ability, enjoyment, and confidence in science. Again, based on previous research, it was hypothesized that racial minority students would report higher rates of bullying victimization than their White peers. Results obtained using the science-related covariates were remarkably similar to those results gathered when using the math-related covariates. As before, the overall model using the covariates of science skills, science enjoyment, and science confidence resulted in a statistically significant difference between the racial demographic groups that demonstrated an increase in the measured difference between the groups and an increased effect size compared to the model without covariates. Hispanic students continued to report significantly higher rates of bullying victimization than White and Multi-Racial students as reported by post hoc analyses. However, the difference between Hispanic and Black students was also shown to be significantly different, as well, with Hispanic students reporting more frequent victimization.

Again, this statistically significant result leads to the conclusion that the null hypothesis should be rejected. However, this result also represents a limited number of differences between the racial groups as well as a very small effect size. While Hispanic students continue to report significantly more frequent rates of bullying victimization, no other racial demographic group reported significantly higher rates of bullying than any other group. Therefore, this significant result should be interpreted with caution.

**Research question 3.** The third research question asked if a student's gender has an impact on his or her experience of peer bullying victimization after controlling for the student's ability, enjoyment, and confidence in mathematics. This ANCOVA procedure resulted in a statistically significant finding, indicating that a student's gender is significantly related to his or her reports of bullying victimization. Female students were shown to report more frequent peer bullying victimization than their male peers. The inclusion of the math-related covariates also resulted in an increase in the difference between the genders in terms of frequency of bullying, and also demonstrated an increase in the model's effect size when compared to the ANOVA calculations completed without the covariates. These results indicate that the null hypothesis should be rejected. Still, the differences found between the genders in bullying victimization resulted in a very small effect size, as well, indicating that the gender differences between bullying victimization frequency have little noticeable impact upon the incidence of being bullied.

**Research question 4.** Finally, the fourth research question asked if a student's gender has an impact on his or her experience of peer bullying victimization after controlling for the student's ability, enjoyment, and confidence in science. Like the other ANCOVA analyses, this model resulted in a statistically significant difference, with a slight increase in the difference found between the groups and an increase in effect size when compared to the previous ANOVA analysis without the science-related covariates. Again, female students reported significantly more frequent experiences of bullying victimization. This result indicates that the null hypothesis for this research question should also be rejected. Again, however, this analysis also reported a very small effect size.

## Conclusions

The results of the completed analyses indicate that there is a statistically significant difference related to the frequency that students who belong to different racial demographic groups report experiencing peer bullying victimization. This significant difference in the reported frequency of peer bullying victimization also exists between male and female students. Furthermore, while this significant difference is shown to exist when both the racial demographic and gender groups are compared directly, the addition of math- and science-related covariates impacted the magnitude of these relationships. After controlling for the effects of students' skill, enjoyment, and confidence in both math and science, a larger difference was found between the racial demographic and gender groups, and this difference was shown to account for a greater portion of the variation in students' reports of peer bullying victimization. With these results, it should be concluded that students do report different frequencies of peer bullying victimization depending on the racial and gender groups that the student belongs to, and that a student's skill, enjoyment, and confidence in math and science are also related to reports of bullying.

While these results did reach a level of statistical significance, the effect sizes generated by the developed models indicated that a student's race and gender explained a very small proportion of the variance in his or her reports of peer bullying victimization. This significant result accompanied by a small effect size indicates that differences found between the racial and gender groups is largely due to the large sample size and high level of statistical power allowed by the TIMSS 2011 dataset. Therefore, the differences between the groups appear trivial and do not reflect a noticeable or practically relevant relationship between the racial and gender groups.

This trivial relationship exists when considering the developed models that do not include any covariates, as well as those models that do include the math- and science-related covariates. Even though the models that included the covariates did explain a greater portion of the variance in peer bullying victimization, the demonstrated increase in effect size remained trivial. This small increase in the covariate models' ability to explain students' responses on the outcome variable indicates that math and science skill, enjoyment, and confidence does little to practically explain students' bullying reports.

**Race-Specific Results**

Even though these main analyses have produced results that make it difficult to draw strong conclusions, the patterns of demonstrated relationships indicated by the main and preliminary analyses are consistent with other results discussed in the literature, and/or with the general hypotheses of this investigation. First, the data presented by the TIMSS 2011 indicate that Hispanic students report significantly higher rates of bullying victimization than White students and Multi-Racial students. This finding is consistent with the generated hypotheses of this study, as racial minority students were thought to experience higher rates of bullying victimization than were their White peers. However, this result adds to the growing inconsistency in the bullying literature regarding the impact of students' race on students' experiences of bullying victimization.

In general, the previous literature has suggested that racial minority students report less frequent experiences of bullying victimization than do their White peers (DeVoe, Peter, Noonan, Snyder, & Baum, 2005; Hanish & Guerra, 2000; Peguero et al., 2011; Williams & Peguero, 2013). However, the specific racial group order related to the reports of bullying victimization varies. In some studies, significant differences are found between White and Hispanic students (DeVoe et al., 2005; Dinkes, Kemp, & Baum, 2009; Hanish & Guerra, 2000), White and Black students (Spriggs, Iannotti, Nansel, & Haynie, 2007; Wang et al., 2009), White and Asian students (Dinkes et al., 2009; Peguero, 2009), Black and Hispanic students (Peskin, Tortolero, & Markham, 2006; Spriggs et al., 2007), and Hispanic and Asian students (Dinkes et al., 2009; Robers et al., 2012).

Across these reviewed studies, it appears as though Hispanic students have, at times, been grouped with White students in terms of reported rates of bullying victimization. This indicates that, often, there are no statistically significant differences between White and Hispanic students' reports of bullying victimization (Robers et al., 2012; Spriggs et al., 2007). The current study does provide some support for the idea that White and Hispanic students experience similar patterns of social repercussions when those students demonstrate academic success. The correlations calculated between students' math and science performance and students' reported frequency of bullying victimization indicate that both White and Hispanic students show statistically significant positive relationships between math performance and rate of bullying victimization. Therefore, for both White and Hispanic students, as a student's math ability increases, that student tends to report higher rates of bullying victimization. No other statistically significant relationship was found between the academic performance and bullying rates for any other racial group. This finding suggests that the ways that peers respond to the academic success demonstrated by White and Hispanic students may be different that the patterns of responses provided to other racial minority groups.

However, one study has found that there was evidence to suggest that Hispanic students' participation in bullying behaviors, both as the perpetrator of bullying and as the victim, was higher than White students when specific types of bullying behaviors were considered (Wang et al., 2009). Finally, the investigations by Peguero and Popp (2012) and by Fryer (2006) indicate that minority students, and Hispanic students in particular, appear to respond differently to the academic success of their peers than do other racial minority groups. These reactions towards students' increased participation in academic-related activities and academic success has been shown to lead to increased episodes of social isolation and peer bullying victimization. The results of this study provide further evidence to support these conclusions that academic participation and success may, in fact, bring about negative social responses for Hispanic students.

**Gender-Specific Results**

The literature related to the rates of bullying victimization reported by male and female students has demonstrated a similar level of inconsistency. Many early studies concluded that male students were more likely to bully their peers and also experience bullying victimization (Björkqvist et al., 1992; Crick et al., 1996). However, as conceptualizations of bullying grew to include more indirect behaviors, the gap between the bullying victimization rates experienced by male and female students narrowed or even reversed. Of the more recent nationwide surveys reviewed, some have reported overall bullying victimization rates to be higher for males (DaVoe et al., 2005; Nansel et al., 2001), while others have indicated that females are more likely to report being bullied (Robers et al., 2012).

While the current study did not measure the percentage of male and female students who experienced peer bullying victimization, the TIMSS 2011 data did indicate that female students reported a frequency of bullying victimization that was significantly higher than male students. Based on this finding, the current study concludes that female students are bullied more often than their male peers. This significant difference increased slightly when the male and female students' math and science skill, enjoyment, and confidence were controlled for through the covariate model.

In addition to the evidence provided by the covariate model, the current study found further evidence that indicates the relationship between academic performance and bullying was different for male versus female students. Related to math ability, female students reported a significant positive correlation with bullying victimization rate, while male students indicated no significant relationship between the two variables. When the correlations between science ability and bullying were examined, female students' responses demonstrated a positive, but statistically non-significant relationship. Male students, however, demonstrated a significant negative relationship between science ability and frequency of peer bullying victimization. These correlations indicate that female students tend to experience more frequent bullying victimization as their academic achievement increases. Male students' results, however, suggest either no association between academic performance and bullying, or less frequent victimization as male students' academic performance improves. Therefore, the present investigation suggests that the way in which academic success impacts the social standing of female students may be qualitatively different from how academic success affects their male peers.

These results are consistent with the theory of "backlash effects" that is discussed by Rudman and Fairchild (2004). These researchers found that female participants were more likely to be sabotaged or punished when they performed in unexpected or counterstereotypic ways. In this study, male and female participants were asked to compete against each other in a trivia competition related to knowledge of traditionally masculine domains (e.g., sports, fighting) or a separate competition related to knowledge of female beauty products. Results indicated that, if a female participant were to win the completion related to the masculine domain, the defeated male participant was less likely to help the female participant succeed in a subsequent competition.

Furthermore, this sabotage was associated with an increase in self-esteem for the defeated male participant. These authors concluded that the findings of this study supported the hypothesis that when an individual's performance does not match gender-stereotypic behavior, that individual is at an increased risk for negative social responses. The findings of the current study appear to support this claim, as the social responses to female students who display academic success appear different, and more negative, than the social responses experienced by male students.

**Bullying of Low Achieving Students**

One of the early assumptions that shaped the perspective of the current study was that academically successful students experience elevated rates of bullying behaviors. While this assumption was grounded in academic research (e.g., Peterson & Ray, 2006), popular culture also influenced this perspective. There are consistent depictions of academic-focused individuals facing social isolation and ridicule. Some clear examples of these story lines include the *Revenge of the Nerds* movie series (Bart & Kanew, 1984) and the *Big Bang Theory* television series (Lorre & Prady, 2007). Throughout these examples, pejorative terms such as "nerd" and "geek" are consistently used in such a way that would qualify as bullying behavior. Given the popularity of these examples, it appears as though the idea that academic success and/or cognitive ability is something to be mocked or leads an individual to be less socially successful is pervasive across American society.

Based on this assumption, the current study attempted to measure if academic success was associated with higher rates of bullying victimization. Results indicate that there is some correlational evidence that supports these hypotheses, as female and racial minority students who demonstrate higher levels of academic performance, enjoyment, and confidence also report more frequent rates of peer bullying victimization. However, while this finding was statistically significant, the practical significance of this result was extremely small. Therefore, it does not appear as though increased academic success is noticeably related to bullying victimization.

Instead, other analyses indicated that the opposite relationship exists when the entire population is considered. When three discrete groups were created based on an increasing frequency of reported bullying victimization, the group that reported the most frequent experience of bullying consistently earned the lowest scores on the Math and Science Assessments. Those students who reported the most frequent bullying victimization also reported the lowest levels of enjoyment in both math and science and the lowest levels of confidence in both math and science.

There is a strong literature base that has concluded that low achieving students are at a greater risk for being bullied. Previous research indicates that students who had been diagnosed with a Specific Learning Disability and placed in an inclusive, rather than a pullout, classroom were more likely to report bullying victimization than their more academically successful peers (Luciano & Savage, 2007). Other research has found that students who were somehow different from their classmates, whether in appearance, academic ability, both at the higher and lower ends, or disability category, were more likely to report bullying victimization (Bishop, 2004; Sweating & West, 2001). Finally, there is evidence to suggest that high levels of academic performance and cognitive abilities are insulating factors against the experience of bullying victimization, as these students tend to possess the social reasoning skills to successfully respond to social stress like bullying (Peguero et al., 2011; Preuss & Dubow, 2004).

Given this literature base, as well as the findings of the current study, it appears as though conventional wisdom that asserts that intellectually gifted and/or academically successful students are at risk for bullying victimization does not accurately portray most students' experiences. Therefore, future research and intervention should continue to target those students who are functioning below academic expectations in order to help support those students' academic skills as well as their social interactions.

**Evidence of the Gender and Racial Achievement Gap**

While the racial and gender differences in the relationship between academic success and bullying victimization remains somewhat unclear, the TIMSS 2011 provides clear evidence that a math and science achievement gap continues to exist between racial and gender groups. As discussed above, this gap in academic achievement has been shown to exist for at least the past 50 years (Coleman et al., 1966), and remains a significant topic of public debate and governmental intervention (United States Government Accountability Office, 2012). However, research indicates that, despite this attention and intervention, both the gender (Barone, 2011) and racial (Aronson et al., 1998) achievement gaps persist. These studies state that male students continue to outperform female students, and White students outperform Black and Hispanic students, in terms of math and science achievement (Aronson et al., 1998). However, one racial minority group, Asian students, has been consistently shown to perform as well as, or better, than White students (Kao & Thompson, 2003). This tendency for Asian students to deviate from the trend of racial minority students earning lower levels of academic performance has led some researchers to conclude that Asian American students are the "model minority" (Kao, 1995).

Results of the Math and Science Assessments administered for TIMSS 2011 match this description of the academic achievement gap perfectly. Male students consistently outperformed female students in both math and science, with gender differences found on the Science Assessment being especially large. Furthermore, White students were shown to earn significantly higher scores on both the Math and Science Assessment than all racial minority groups except for Asian students. When compared to Asian students, White students were shown to earn significantly lower scores on the Math Assessment, while the two groups' scores were not shown to be significantly different on the Science Assessment. This appears to provide further evidence for Asian students' classification as the "model minority."

Not only are there significant differences in actual math and science performance, but also there appear to be significant gender and racial differences in terms of student-reported levels of enjoyment and confidence in those academic subjects. These differences also tend to mirror the differences found in actual academic performance; however, with some interesting differences. When gender differences were calculated, male students were shown to report significantly higher levels of enjoyment in both math and science, as well as significantly higher levels of confidence in both math and science. When racial differences were examined, Asian students reported the highest levels of enjoyment and confidence in math and science. However, while White students reported higher levels of confidence in both math and science, as well as higher levels of enjoyment in science than most other racial minority groups, both Black and Hispanic students reported higher levels of enjoyment in math than did their White peers.

These results indicate that, in the sample studied by the TIMSS 2011, female and racial minority students who perform well on the Math and Science Assessments would still demonstrate a noticeable deviation from their respective group means. Therefore, it is reasonable to conclude that the gender and racial stereotypes that have traditionally been associated with academic performance (Fryer, 2006; Ogbu, 2003; Renold & Allen, 2006) are likely to persist to the present day. However, the results obtained by the current study related to the impact that academic performance, enjoyment, and confidence has on bullying victimization does little to practically explain the maintenance of these performance gaps and the gender- and racial-based stereotypes. In order to better explain how and why these gaps persist, further research is needed.

## Limitations

As with most studies that utilize secondary data analysis, the most important limitation is related to the questions posed in the original data-collection measures, and the validity of those questions relating to the current study's research questions. On the TIMSS 2011 Student Questionnaire, the five questions on the Student Bullied at School Scale did appear to be strongly related to the research-based definitions of verbal aggression (Camodeca & Goosens, 2005), social aggression (Archer & Coyne, 2005), relational aggression (Cullerton-Sen & Crick, 2005), and physical aggression (Card et al., 2008). Given the questions' close associations with research-based definitions of bullying, these questions appear valid and able to accurately measure bullying behaviors. However, the small number of questions related to each type of aggressive behavior limits the reliability of the results related to each distinct bullying behavior. Therefore, in order to increase the reliability of each student's bullying results, the entire Student Bullied at School Scale score was used as the dependent variable.

By using the Scale variable, however, new data- and interpretation-based limitations arise. First, when using the Scale variable, there is a limited range of score possibilities on the scale because due to the fact that it consists of only five items, with each item containing only four possible responses. Second, bullying research has consistently shown that the different types of aggressive behaviors that can be used to bully others do not occur at the same rates (Nansel et al., 2001; Robers et al., 2012; Wang et al., 2009). Therefore, it is somewhat unlikely that a student who experiences one type of bullying victimization will be just as likely to experience a different type of bullying victimization as well. This means, for example, that a student who experiences relational aggression daily at school, and answered appropriately on the relational aggression-related question, but does not experience elevated rates of physical, verbal, and social aggression will demonstrate a somewhat low overall score on the Student Bullied at School Scale. Furthermore, because two of the questions on the five question scale are related to physical bullying, those students who do experience physical bullying are likely to earn significantly higher scores on the scale than are students who are bullied just as frequently, but experience different types of aggression. These features of the Student Bullied at School Scale complicate the interpretation of the gathered findings.

Another limitation when relying on bullying frequency data gathered using brief survey techniques, like the TIMSS 2011, is that it is generally impossible to gauge the characteristics of the specific aggressive student or students who is responsible for bullying the individual who is reporting the bullying experience. This lack of complete understanding regarding the bullying relationship makes it difficult to determine if the experiences reported do, in fact, meet the tripartite definition of bullying (Gottheil & Dubow, 2001). In order to be considered an episode of bullying, a power differential must exist between the aggressor and the victim. Given that this concept of power differential is somewhat broad, and can include such qualities as size, socio-economic status, peer group standing, and cognitive ability (Sutton et al., 1999), it could be argued that any student might demonstrate at least one type of power differential over any other student. However, this is still impossible to gauge with certainty based on the information gathered by the TIMSS 2011.

Likewise, it is equally unclear if the behaviors students identify as bullying also meet the qualification of being proactive or instrumental aggression. Students may report any type of aggressive behavior, such as fights that occur in response to a provocation, as an example of bullying, even though the behavior was not a premeditated or goal-oriented aggressive act. Ultimately, researchers who gather this type of data must rely on the face validity of the questions and on the ability of students to understand the difference between bullying and simple aggression; even when there is research to suggest that students are, at times, unable to successfully make this distinction (Monks & Smith, 2006; Smith et al., 2002).

Another limitation of the current study is related to the generalizability of the findings. While the TIMSS 2011 developers and data collectors conducted a thorough sampling procedure that resulted in a broad sample that was closely related to the characteristics of the student population at large (Kastberg et al., 2013), the current study only examined the results from the 8$^{th}$ grade sample. Students' age is frequently related to reported rates of bullying victimization (Nansel et al., 2001; Robers et al., 2012) and by examining only a single age group, it is likely inappropriate to generalize these results to students of different ages.

Finally, the interpretability of the current findings was also impacted by the data not meeting the assumptions of the ANCOVA procedure. Even though ANCOVA has been shown to be robust to violations of the assumptions (Porter & Raudenbush, 1987), some assumption violations do impact the ability to accurately conclude the impact of the independent variables, gender and racial demographic, on the dependent variable of peer bullying victimization. Most importantly, because the data violated the independence of covariates and independent variables assumption (Field, 2013), the covariates and the independent variables used were shown to account for a shared portion of the variance in the outcome variable. This was true each of the four models examined, and the effect was clearly evidenced by the small increases in effect size of the models that included the covariates in comparison to the models that did not include the covariates. Due to the significant relationship between the covariates and the independent variables, the total unique amount of variance that the covariates explained in the frequency of students' bullying experiences is difficult to determine.

**Future Directions**

Results gathered from the current study give rise to other questions that further research could address; several of which could involve further examination of the TIMSS 2011 dataset to provide valuable insights. First, the TIMSS 2011 dataset could provide interesting information related to the frequency of each specific type of aggressive behavior students report, and if each type of bullying behavior impacts academic participation differently. As discussed above, the bullying-related questions included in the TIMSS 2011 Student Survey are closely related to the developed definitions of verbal aggression (Camodeca & Goosens, 2005), social aggression (Archer & Coyne, 2005), relational aggression (Cullerton-Sen & Crick, 2005), and physical aggression (Card et al., 2008). Due to the empirical support of the questions asked, the important information provided by each question trumps the increased reliability provided by using the compiled Student Bullied at School Scale.

As discussed throughout the current study, previous research has described a complicated, and somewhat inconsistent, relationship between bullying experiences and academic performance. However, the majority of the studies examined have either investigated the long-term academic effects of bullying overall, without examining the specific types of aggressive behaviors, or the research has simply measured the overall rates of each type of aggressive behavior without comparing the long-term academic outcomes of students who experience each type of aggression. For example, one longitudinal research studied concluded that young children who displayed stable levels of aggressive behavior across several years of early development tended to demonstrate lower levels of academic achievement at age 12 (Campbell, Spieker, Burchinal, & Poe, 2006). However, this research did not examine the effects that each specific type of aggressive behavior had on students' academic performance.

One study that did begin to investigate outcomes of each type of aggressive behavior found that physical aggression, specifically, was significantly related to lower grade point averages (GPA) in high school students, and that physical aggression added predictive power of students' GPA above students' measured personality traits (Loveland, Lounsbury, Welsh, & Bublotz, 2007). This study went on to conclude that this relationship between academic functioning and physically aggressive behavior was different for male and female students (Loveland et al., 2007). While this is an important finding related to physical aggression, more research is needed to determine the specific impact of verbal, relational, and social aggression, as well. This gap in the literature, which has been noted by other researchers (Crick, Ostrov, & Werner, 2006), appears to be an area where research, in general, and the TIMSS 2011 dataset, specifically, could shed light.

Another area future research could continue to examine is ways in which to increase female students' enjoyment and confidence in math and science coursework. The current study determined that female students continue to report lower levels of enjoyment and confidence in both math and science coursework than do their male peers. This significant difference has been shown to exist for some time (e.g., Eccles, 1989), and does not seem to have diminished. Recent investigations have found that this discrepancy between male and female students' perceptions of math and science is related to the number of students enrolled in a class (Sobel, Gilmartin, & Sankar, 2016), the perceptions and implicit expectations of female students' parents and teachers (Gunderson, Ramirez, Levine, & Beilock, 2012), and the social supportiveness of female students' peers in relation to their involvement in math and science courses (Rice, Barth, Guadagno, Smith, & McCallum, 2013). These and other research studies provide insights into the areas where further intervention would be beneficial.

Given the consistency of the confidence and enjoyment gender gaps, however, the most needed research appears related to specific interventions and programs that would directly combat the gaps. One such research-based program is the Bringing Up Girls in Science (BUGS) afterschool program (Tyler-Wood, Ellison, Lim & Periathiruvadi, 2012). The BUGS program is a female mentoring program for 4$^{th}$ and 5$^{th}$ grade female students who are exposed to environmental science concepts by female professional scientists. After completing the BUGS program, those girls who participated demonstrated significant growth in science knowledge, more positive perceptions of science as a whole, and more positive perceptions of future science-related careers than did their peers who did not participate in the BUGS intervention (Tyler-Wood et al., 2012). Further research and implementation of programs whose purpose is similar to the BUGS program appears to be vital to increasing female students' participation and success in math and science.

Further research is also warranted into the racial differences of bullying experiences. One interesting finding of the current study is that Hispanic students reported significantly higher rates of bullying victimization than their White peers. This finding appears to contradict previous research that has found that Hispanic students report less frequent rates of bullying victimization than their White peers (Hanish & Guerra, 2000; Peguero et al., 2011; Williams & Peguero, 2013). However, the United States is currently in the midst of a significant shift in the populations' racial demographics. The United States Census Bureau reports that White Americans represented a substantial racial majority in the year 2014, with White Americans comprising 62.2% of the population. By the year 2060, however, and likely before, the United States population will no longer possess a single racial majority group, as White Americans will account for only 43.6% of the population, while Hispanic Americans' population will continue to rise and will remain the second most populous racial group, accounting for roughly 28.6% of the total population (Colby & Ortman, 2015).

While a White majority persists in the population at large today, enrollment in the United States' educational system indicates that this plurality of racial demographics has already taken place. In the fall of 2014, racial minority students were shown to account for 50.3% of the total enrollment in American schools (Maxwell, 2014). Furthermore, by the year 2060, the U.S. Census Bureau projects that the population of individuals living in the United States who are under the age of 18 will consist of 35.6% White students and 33.5% Hispanic students (Colby & Ortman, 2015). This shift in schools' racial demographics may impact how students relate to each other; including the reported experiences of bullying victimization. The bullying victimization data presented by the TIMSS 2011 may be one of the first indications of how this shift may manifest.

Finally, the last recommendation for future research is related to how students respond to being bullied by others who are a part of the same racial group. While the current study examined each racial groups' reported rates of bullying victimization, one of the stated limitations of the study is that it is difficult to know the characteristics of the student who is responsible for perpetrating the bullying behaviors. However, previous literature suggests that, especially when examining racial bullying differences, it is important to consider the racial characteristics of the bully as well as the victim.

Some research suggests that when students are more likely to be bullied by other members of their same racial group. For example, in their research related to the "acting White" phenomenon, Fordham and Ogbu (1986) and Ogbu (2003) concluded that Black students who violated racial stereotypes were at an increased risk to experience bullying victimization from other Black students in an attempt to maintain group conformity. Some studies have also found that when students report being bullying by students of the same racial group, those bullying victims report more significant negative outcomes including loneliness and social anxiety (Bellmore et al., 2004; Graham et al., 2009).

Other research, however, has found that grouping racially similar students together is associated with more positive bullying and academic related outcomes. One study found that when bullying rates reported by students who were enrolled in racially homogeneous classroom were compared to the bullying experiences reported by students enrolled in racially heterogeneous classrooms, those student who attended racially heterogeneous classrooms reported higher rates of bullying victimization (Vervoort, Scholte, & Overbeek, 2010). Furthermore, one nationwide survey of students found that when students attend highly racially segregated schools, including schools that enroll a majority of Black and Latino students, those students report a more optimistic outlook on academic achievement and have a more positive perspective about school (Goldsmith, 2004).

Again, the TIMSS 2011 dataset may help answer this question. Due to the meticulous sampling procedure, the racial compositions of the classrooms studied could be calculated, and then the bullying responses and academic performances of the students in those classrooms could be compared. The findings, then, may provide further insights into what types of student groups may need the most intensive interventions related to bullying perpetration and academic perceptions.

## Summary

The current investigation found that students from different gender and racial groups did report significantly different rates of bullying victimization after controlling for those students' math and science abilities, enjoyment, and confidence. However, these differences did not explain a large amount of the variance in students' bullying victimization reports. Therefore, the current models indicate that math and science ability, enjoyment, and confidence are not the most important factors related to students' bullying experiences.

While the results of the main analyses were less meaningful than hypothesized, other preliminary analyses did produce some interesting results. First, significant differences were found in the rates at which different gender and racial groups report experiencing bullying victimization, with female students reporting more frequent victimization than male students, and Hispanic students reporting more frequent victimization than White and Multi-Racial students. It also appears as though gender, specifically, plays a noticeable role in the way students' bullying experiences relate to their academic achievement. Female students were shown to demonstrate more frequent bullying as their math and science performance increased, while male students showed no relationship between math performance and bullying, and a negative relationship between science performance and bullying victimization. Second, contrary to the main assumptions of the current study, there is evidence to suggest that, in general, those students who report experiencing the most frequent rates of bullying victimization also earn the lowest scores on the math and science assessments, and also report the lowest levels of enjoyment and confidence in both math and science. Third, the TIMSS 2011 data indicate that the math and science academic achievement gaps between gender and racial minority groups continue to persist.

Limitations to the current study include concerns related to the overall validity of the data related to bullying victimization. The primary concern is due to the nature of secondary data analysis, as the questions asked of students were not specifically geared toward the completion of the present investigation. It is also unclear if students' reports do reflect experiences that match each of the three criteria for bullying behaviors. Another limitation with the current study is related to the selection of the ANCOVA statistical technique, as the TIMSS 2011 violated several assumptions of this procedure, which impacted the interpretability of the main analyses.

Finally, ideas for the directions for future research spawned by the current study include investigations into the rates of each specific type of aggressive action that is discussed in the TIMSS 2011 questionnaire.  Another area of needed research is into how to increase female students' enjoyment and confidence in math and science; in hopes that improvement in these areas will also help improve female students' overall performance in math and science.  The last suggestion for further investigation focuses on the racial differences in reports of bullying victimization, and how intra-group victimization impacts students who identify as a racial minority.

## REFERENCES


Archer, J., & Coyne, S. M. (2005). An integrated review of indirect, relational, and social aggression. *Personality and Social Psychology Review*, *9*, 212–230. doi:10.1207/s15327957pspr0903_2

Aronson, J., Quinn, D. M., & Spencer, S. J. (1998). Stereotype threat and the academic underperformance of minorities and women. In J. K. Swim & C. Stangor (Eds.), *Prejudice: The target's perspective* (pp. 83–103). San Diego, CA, US: Academic Press.

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, *84*, 191–215. doi:10.1037/0033-295X.84.2.191

Barone, C. (2011). Some things never change: Gender segregation in higher education across eight nations and three decades. *Sociology of Education*, *84*, 157–176. doi:10.1177/0038040711402099

Bart, P. (Producer), & Kanew, J. (Director). (1984). *Revenge of the nerds* [Motion Picture]. USA: Twentieth Century Fox Film Corporation.

Beaton, A. E., & González, E. (1995) *The NAEP primer*. Chestnut Hill, MA: Boston College.

Bellmore, A. D., Witkow, M. R., Graham, S., & Juvonen, J. (2004). Beyond the individual: The impact of ethnic context and classroom behavioral norms on victims' adjustment. *Developmental Psychology*, *40*, 1159–1172. doi:10.1037/0012-1649.40.6.1159

Beran, T., & Li, Q. (2007). The relationship between cyberbullying and school bullying. *Journal of Student Wellbeing*, *1*, 15–33.

Bishop, J. H., Bishop, M., Bishop, M., Gelbwasser, L., Green, S., Peterson, E., … Zuckerman, A. (2004). Why we harass nerds and freaks: A formal theory of student culture and norms. *Journal of School Health*, *74*, 235–251. doi:10.1111/j.1746-1561.2004.tb08280.x

Björkqvist, K., Lagerspetz, M. J., & Kaukiainen, A. (1992). Do girls manipulate and boys fight? Developmental trends in regard to direct and indirect aggression. *Aggressive Behavior*, *18*, 117–127. doi:10.1002/1098-2337(1992)18:2<117::AID-AB2480180205>3.0.CO;2-3

Booth, W. C., Colomb, G. G., & Williams, J. M. (2008). *The craft of research* (3rd ed.). Chicago, IL: University of Chicago Press.

Bradshaw, C. P., Sawyer, A. L., & O'Brennan, L. M. (2007). Bullying and peer victimization at school: Perceptual differences between students and school staff. *School Psychology Review*, *36*, 361–382.

Bronfenbrenner, U. (1977). Toward an experimental ecology of human development. *American Psychologist*, *32*, 513–531. doi:10.1037/0003-066X.32.7.513

Brown, B. B. & Steinberg, L. (1990). Academic achievement and social acceptance. *The Education Digest*, *55*, 57-60.

Brown, K. S. & Parsons, R. D. (1998). Accurate identification of childhood aggression: A key to successful intervention. *Professional School Counseling*, *2*, 135–140.

Camodeca, M., & Goossens, F. A. (2005). Aggression, social cognitions, anger, and sadness in bullies and victims. *Journal of Child Psychology and Psychiatry*, *46*, 186–197. doi:10.1111/j.1469-7610.2004.00347.x

Campbell, S. B., Spieker, S., Burchinal, M., & Poe, M. D. (2006). Trajectories of aggression from toddlerhood to age 9 predict academic and social functioning through age 12. *Journal of Child Psychology and Psychiatry*, *47*, 791–800. doi:10.1111/j.1469-7610.2006.01636.x

Card, N. A., & Hodges, E. V. E. (2008). Peer victimization among schoolchildren: Correlations, causes, consequences, and considerations in assessment and intervention. *School Psychology Quarterly*, *23*, 451–461. doi:10.1037/a0012769

Card, N. A., Stucky, B. D., Sawalani, G. M., & Little, T. D. (2008). Direct and indirect aggression during childhood and adolescence: A meta-analytic review of gender differences, intercorrelations, and relations to maladjustment. *Child Development*, *79*, 1185–1229. doi:10.1111/j.1467-8624.2008.01184.x

Chen, X., & Weko, T. (2009). *Students who study science, technology, engineering, and mathematics (STEM) in postsecondary education* (No. NCES 2009-161) (pp. 1–24). Washington D.C.: U.S. Department of Education.

Cohen, G. L., Garcia, J., Apfel, N., & Master, A. (2006). Reducing the racial achievement gap: A social-psychological intervention. *Science*, *313*, 1307–1310. doi:10.1126/science.1128317

Colby, S. L., & Ortman, J. M. (2015). *Projections of the size and composition of the U.S. populaiton: 2014 to 2060* (Current Population Reports No. P25-1143). Washington D.C.: U.S. Census Bureau.

Coleman, J. S., Campbell, E. Q., Hobson, C. J., McPartland, J., Mood, A. M., Weinfeld, F. D., & York, R. I. (1966). *Equality of educational opportunity* (No. OE-38001) (pp. 1–34). Washington D.C.: U.S. Department of Health, Education, and Welfare.

Cook, C. R., Williams, K. R., Guerra, N. G., Kim, T. E., & Sadek, S. (2010). Predictors of bullying and victimization in childhood and adolescence: A meta-analytic investigation. *School Psychology Quarterly*, *25*, 65–83. doi:10.1037/a0020149

Coyne, S. M., Archer, J., & Eslea, M. (2006). “We’re not friends anymore! Unless...”: The frequency and harmfulness of indirect, relational, and social aggression. *Aggressive Behavior*, *32*, 294–307. doi:10.1002/ab.20126

Craig, W. M. (1998). The relationship among bullying, victimization, depression, anxiety, and aggression in elementary school children. *Personality and Individual Differences*, *24*, 123–130. doi:10.1016/S0191-8869(97)00145-1

Crick, N. R., Bigbee, M. A., & Howes, C. (1996). Gender differences in children’s normative beliefs about aggression: How do I hurt thee? Let me count the ways. *Child Development*, *67*, 1003–1014. doi:10.1111/j.1467-8624.1996.tb01779.x

Crick, N. R., Casas, J. F., & Ku, H. C. (1999). Relational and physical forms of peer victimization in preschool. *Developmental Psychology*, *35*, 276–385. doi:10.1111/j.1467-8624.1996.tb01779.x

Crick, N. R., & Grotpeter, J. K. (1995). Relational aggression, gender, and social-psychological adjustment. *Child Development*, *66*, 710–722. doi:10.1111/j.1467-8624.1995.tb00900.x

Crick, N. R., Ostrov, J. M., & Werner, N. E. (2006). A longitudinal study of relational aggression, physical aggression, and children’s social–psychological adjustment. *Journal of Abnormal Child Psychology*, *34*, 127–138. doi:10.1007/s10802-005-9009-4

Crothers, L. M., Schreiber, J. B., Field, J. E., & Kolbert, J. B. (2009). Development and measurement through confirmatory factor analysis of the Young Adult Social Behavior Scale (YASB): An assessment of relational aggression in adolescence and young adulthood. *Journal of Psychoeducational Assessment*, *27*, 17–28. doi:10.1177/0734282908319664

Crowley, M. F. (1977). *Women and Minorities in Science and Engineering*. Washington, D.C.: National Science Foundation. Retrieved from http://eric.ed.gov/?id=ED147098

Cullerten-Sen, C., & Crick, N. R. (2005). Understanding the effects of physical and relational victimization: The utility of multiple perspectives in predicting social-emotional adjustment. *School Psychology Review*, *34*, 147–160.

Dalton, B., Ingels, S. J., Downing, J., & Bozick, R. (2007). Advanced mathematics and science coursetaking in the spring high school senior classes of 1982, 1992, and 2004 (NCES 2007-312). Washington, DC: National Center for Education Statistics. Retrieved from http://eric.ed.gov/?id=ED497754

DaVoe, J. F., Peter, K., Noonan, M., Snyder, T. D., & Baum, K. (2005). *Indicators of School Crime and Safety: 2005* (No. NCJ 210697). Washington D.C.: U.S. Department of Education.

Deary, I. J., Penke, L., & Johnson, W. (2010). The neuroscience of human intelligence differences. *Nature Reviews Neuroscience*, *11*, 201–11. doi:10.1038/nrn2793

Dinkes, R., Kemp, J., & Baum, K. (2009). *Indicators of School Crime and Safety: 2008* (No. NCJ 226343). Washington D.C.: U.S. Department of Education.

Eccles, J. S. (1989). Bringing young women to math and science. In M. Crawford & M.

Gentry (Eds.), *Gender and Thought* (pp. 36–57). New York, NY: Springer-Verlag.

Erdur-Baker, Ö. (2010). Cyberbullying and its correlation to traditional bullying, gender and frequent and risky usage of internet-mediated communication tools. *New Media & Society*, *12*(1), 109–125. doi:10.1177/1461444809341260

Espelage, D. L., & Swearer, S. M. (2003). Research on school bullying and victimization: What have we learned and where do we go from here? *School Psychology Review*, *32*, 365–383.

Estell, D. B., Farmer, T. W., Irvin, M. J., Crowther, A., Akos, P., & Boudah, D. J. (2009). Students with exceptionalities and the peer group context pf bullying and victimization in late elementary school. *Journal of Family Studies*, *18*, 136–150. doi:10.1007/s10826-008-9214-1

Fekkes, M., Pijpers, F. I. M., Fredriks, A. M., Vogels, T., & Verloove-Vanhorick, S. P. (2006). Do bullied children get ill, or do ill children get bullied? A prospective cohort study on the relationship between bullying and health-related symptoms. *Pediatrics*, *117*, 1568–1574. doi:10.1542/peds.2005-0187

Field, A. (2013). *Discovering statistics using IBM SPSS statistics* (4th ed.). Thousand Oaks, CA: Sage.

Field, J. E., Kolbert, J. B., Crothers, L. M., & Hughes, T. L. (2009). *Understanding girl bullying and what to do about it: Strategies to help heal the divide*. Thousand Oaks, CA: Corwin.

Fonagy, P., Twemlow, S. W., Vernberg, E., Sacco, F. C., & Little, T. D. (2005). Creating a peaceful school learning environment: The impact of an antibullying program on educational attainment in elementary schools. *Medical Science Monitor*, *11*, 317–325.

Fordham, S., & Ogbu, J. U. (1986). Black students' school success: Coping with the "burden of 'acting white.'" *The Urban Review*, *18*, 176–206. doi:10.1007/BF01112192

Fryer, R. G. (2006). "Acting white": The social price paid by the best and brightest minority students. *Education Next*, *6*, 53–59.

Galen, B. R., & Underwood, M. K. (1997). A developmental investigation of social aggression among children. *Developmental Psychology*, *33*, 589–600. doi:10.1037/0012-1649.33.4.589

Glass, G. V. (1976). Primary, secondary, and meta-analysis of research. *Educational Researcher*, *5*, 3–8. doi:10.3102/0013189X005010003

Glew, G. M., Fan, M.-Y., Katon, W., & Rivara, F. P. (2008). Bullying and school safety. *The Journal of Pediatrics*, *152*, 123–128. doi:10.1016/j.jpeds.2007.05.045

Goldsmith, P. A. (2004). Schools' racial mix, students' optimism, and the Black-White and Latino-White achievement gaps. *Sociology of Education*, *77*, 121–147. doi:10.1177/003804070407700202

Good, C., Aronson, J., & Harder, J. A. (2008). Problems in the pipeline: Stereotype threat and women's achievement in high-level math courses. *Journal of Applied Developmental Psychology*, *29*, 17–28. doi:10.1016/j.appdev.2007.10.004

Gottheil, N. F., & Dubow, E. F. (2001). Tripartite belief models of bully and victim behavior. *Journal of Emotional Abuse, 2,* 25-47. doi:10.1300/J135v02n02_03

Gould, S. J. (1981). *The Mismeasure of Man*. New York, NY: Norton & Co.

Graham, S., Bellmore, A., Nishina, A., & Juvonen, J. (2009). “It must be me”: Ethnic diversity and attributions for peer victimization in middle school. *Journal of Youth and Adolescence*, *38*, 487–499. doi:10.1007/s10964-008-9386-4

Gravetter, F., & Forzano, L. A. (2009). *Research methods for the behavioral sciences* (3rd ed.). Belmont, CA: Wadsworth Cengage Learning.

Griffin, R. S., & Gross, A. M. (2004). Childhood bullying: Current empirical findings and future directions for research. *Aggression and Violent Behavior*, *9*, 379–400. doi:10.1016/S1359-1789(03)00033-8

Guerin, S., & Hennessy, E. (2002). Pupils’ definitions of bullying. *European Journal of Psychology of Education*, *17*, 249–261. doi:10.1007/BF03173535

Gunderson, E. A., Ramirez, G., Levine, S. C., & Beilock, S. L. (2012). The role of parents and teachers in the development of gender-related math attitudes. *Sex Roles*, *66*, 153–166. doi:10.1007/s11199-011-9996-2

Hanish, L. D., & Guerra, N. G. (2000). The roles of ethnicity and school context in predicting children’s victimization by peers. *American Journal of Community Psychology*, *28*, 201–23. doi:10.1023/A:1005187201519

Horowitz, M. J. (1979). *Stress response syndromes*. New York, NY: Jason Aronson.

Houbre, B., Tarquinio, C., Thuillier, I., & Hergott, E. (2006). Bullying among students and its consequences on health. *European Journal of Psychology of Education*, *21*, 183–208. doi:10.1007/BF03173576

Hyde, J. S., Lindberg, S. M., Linn, M. C., Ellis, A. B., & Williams, C. C. (2008). Gender similarities characterize math performance. *Science*, *321*, 494–495. doi:10.1126/science.1160364

Janke, R. W., & Lee, K. (1991). Social skills ratings of exceptional students. *Journal of Psychoeducational Assessment*, *9*, 54–66. doi:10.1177/073428299100900105

Juvonen, J., Nishina, A., & Graham, S. (2000). Peer harassment, psychological adjustment, and school functioning in early adolescence. *Journal of Educational Psychology*, *92*, 349–359. doi:10.1037/0022-0663.92.2.349

Kao, G. (1995). Asian Americans as model minorities? A look at their academic performance. *American Journal of Education*, *103*, 121–159. doi:10.1086/444094

Kao, G., & Thompson, J. S. (2003). Racial and ethnic stratification in educational achievement and attainment. *Annual Review of Sociology*, *29*, 417–442. doi:10.1146annurev.soc.29.010202.100019

Kastberg, D., Roey, S., Ferraro, D., Lemanski, N., & Erberber, E. (2013). *U.S. TIMSS and PIRLS 2011 technical report and user's guide* (No. NCES 2013-046). Washington D.C.: National Center for Education Statistics.

Klomek, A. B., Marrocco, F., Kleinman, M., Schonfeld, I. S., & Gould, M. S. (2008). Peer victimization, depression, and suicidiality in adolescents. *Suicide and Life-Threatening Behavior*, *38*, 166–180. doi:10.1521/suli.2008.38.2.166

Kluwin, T. N., & Morris, C. S. (2006). Lost in a giant database: The potentials and pitfalls of secondary analysis for deaf education. *American Annals of the Deaf*, *151*, 121–128.

Kochenderfer, B. J., & Ladd, G. W. (1996). Peer victimization: Cause or consequence of school maladjustment? *Child Development*, *67*, 1305–1317. doi:10.1111/1467-8624.ep9704040996

Ladson-Billings, G. (2006). From the achievement gap to the education debt: Understanding achievement in U.S. schools. *Educational Researcher*, *35*, 3–12. doi:10.3102/0013189X035007003

Lagerspetz, K. M. J., Björkqvist, K., & Peltonen, T. (1988). Is indirect aggression typical of females? Gender differences in aggressiveness in 11- to 12-year-old children. *Aggressive Behavior*, *14*, 403–414. doi:10.1002/1098-2337(1988)14:6<403::AID-AB2480140602>3.0.CO;2-D

Lee, J. (2002). Racial and ethnic achievement gap trends: Reversing the progress toward equity? *Educational Researcher*, *31*, 3–12. doi:10.3102/0013189X031001003

Lorre, C. & Pardy, B. (Producers). (2007). *The big bang theory* [Television series]. Burbank, CA: Columbia Broadcasting System.

Li, Q. (2006). Cyberbullying in schools a research of gender differences. *School Psychology International*, *27*, 157–170. doi:10.1177/0143034306064547

Li, Q. (2010). Cyberbullying in high schools: A study of students’ behaviors and beliefs about this new phenomenon. *Journal of Aggression, Maltreatment, & Trauma*, *19*, 372–392. doi:10.1080/10926771003788979

Lipinski, J., & Crothers, L. M. (2014). *Bullying in the workplace: Causes symptoms, and remedies*. New York, NY: Routledge.

Loveland, J. M., Lounsbury, J. W., Welsh, D., & Buboltz, W. C. (2007). The validity of physical aggression in predicting adolescent academic performance. *British Journal of Educational Psychology*, *77*, 167–176. doi:10.1348/000709905X79563

Luciano, S., & Savage, R. S. (2007). Bullying risk in children with learning difficulties in inclusive educational settings. *Canadian Journal of School Psychology*, *22*, 14–31. doi:10.1177/0829573507301039

Martin, M. O. & Mullis, I. V. S. (Eds). (2011). *TIMSS 2011 field test of mathematics and science items*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College.

Martin, M. O., Mullis, I. V. S., Foy, P., & Arora, A. (2011). *Creating and interpreting the TIMSS and PIRLS 2011 Context Questionnaire Scales*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College.

Maxwell, L. A. (2014, August 20). U.S. school enrollment hits majority-minority milestone. *Education Week*. Retrieved from http://www.edweek.org/ew/articles/2014/08/20/01demographics.h34.html

Mertler, C. A., & Vannatta, R. A. (2009). *Advanced and multivariate statistical methods* (4th ed.). Glendale, CA: Pyrczak.

Monks, C. P., & Smith, P. K. (2006). Definitions of bullying: Age differences in understanding of the term, and the role of experience. *British Journal of Developmental Psychology*, *24*, 801–821. doi:10.1348/026151005X82352

Mueller, C. E. & Hart, C. O. (2010). Effective use of secondary data analysis in gifted education research: Opportunities and challenges. *Gifted Children*, *4*, 6–11.

Mullis, I. V. S. & Martin, M. O. (2011). *TIMSS 2011 item writing guidelines*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College.

Mullis, I. V. S., Martin, M. O., Ruddock, G. J., O'Sullivan, C. Y., & Preuschoff, C. (2009). *TIMSS 2011: Assessment frameworks*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center Lynch School of Education, Boston College.

Nansel, T. R., Overpeck, M., Pilla, R. S., Ruan, W. J., Simons-Morton, B., & Scheidt, P. (2001). Bullying behaviors among US youth: Prevalence and association with psychosocial adjustment. *Journal of the American Medical Association*, *285*, 2094–2100. doi:10.1001/jama.285.16.2094

National Math & Science Initiative. (2013). *The STEM crisis*. Retrieved from http://nms.org/Education/TheSTEMCrisis.aspx

Oakes, J. (1990). Opportunities, achievement, and choice: Women and minority students in science and mathematics. *Review of Research in Education*, *16*, 153–222. doi:10.2307/1167352

Office of Management and Budget. (2013). *Educating a competitive workforce*. Retrieved from http://www.whitehouse.gov/omb/budget/factsheet/educating-a-competitive-workforce

Ogbu, J. U. (2003). *Black American students in an affluent suburb: A study of academic disengagement*. Mahwah, NJ: Lawrence Erlbaum Associates.

Patterson, M. M., & Bigler, R. S. (2007). Effects of physical atypicality on children's social identities and intergroup attitudes. *International Journal of Behavioral Development*, *31*, 433–444. doi:10.1177/0165025407081472

Peguero, A. A. (2009). Opportunity, involvement, and student exposure to school violence. *Youth Violence and Juvenile Justice*, *7*, 299–312. doi:10.1177/1541204008330634

Peguero, A. A., & Popp, A. M. (2012). Youth violence at school and the intersection of gender, race, and ethnicity. *Journal of Criminal Justice*, *40*, 1–9. doi:10.1016/j.jcrimjus.2011.11.005

Peguero, A. A., Popp, A. M., & Koo, D. J. (2011). Race, ethnicity, and school-based adolescent victimization. *Crime & Delinquency*. Advance online publication. doi:10.1177/0011128711398021

Peguero, A. A., & Williams, L. M. (2011). Racial and ethnic stereotypes and bullying victimization. *Youth & Society, 45*, 545-564. doi:10.1177/0044118X11424757

Peskin, M. F., Tortolero, S. R., & Markham, C. M. (2006). Bullying and victimization among Black and Hispanic adolescents. *Adolescence*, *41*, 467–484.

Peterson, J. S., & Ray, K. E. (2006). Bullying and the gifted: Victims, perpetrators, prevalence, and effects. *Gifted Child Quarterly*, *50*, 148–168. doi:10.1177/001698620605000206

Phelan, J. E., & Rudman, L. A. (2010). Reactions to ethnic deviance: The role of backlash in racial stereotype maintenance. *Journal of Personality and Social Psychology*, *99*, 265–281. doi:10.1037/a0018304

Plunker, J. A., Burroughs, N., & Song, R. (2010). *Mind the (other) gap! The growing excellence gap in K-12 education*. Bloomington, IN: Center for Evaluation and Education Policy.

Popp, A. M., & Peguero, A. A. (2011). Routine activities and victimization at school: The significance of gender. *Journal of Interpersonal Violence*, *26*, 2413–2436. doi:10.1177/0886260510383021

Porter, A. C., & Raudenbush, S. W. (1987). Analysis of covariance: Its model and use in psychological research. *Journal of Counseling Psychology*, *34*, 383–392. doi:10.1037/0022-0167.34.4.383

Prentice, D. A., & Carranza, E. (2002). What women and men should be, shouldn't be, are allowed to be, and don't have to be: The contents of prescriptive gender stereotypes. *Psychology of Women Quarterly*, *26*, 269–281. doi:10.1111/1471-6402.t01-1-00066

Preuss, L. J., & Dubow, E. F. (2004). A comparison between intellectually gifted and typical children in their coping responses to a school and a peer stressor. *Roeper Review*, *26*, 105–111. doi:10.1080/02783190409554250

Renold, E., & Allan, A. (2006). Bright and beautiful: High achieving girls, ambivalent femininities, and the feminization of success in the primary school. *Discourse: Studies in the Cultural Politics of Education*, *27*, 457–473. doi:10.1080/01596300600988606

Rice, L., Barth, J. M., Guadagno, R. E., Smith, G. P. A., & McCallum, D. M. (2013). The role of social support in students' perceived abilities and attitudes toward math and science. *Journal of Youth and Adolescence*, *42*, 1028–1040. doi:10.1007/s10964-012-9801-8

Richardson, D. S., & Green, L. R. (2006). Direct and indirect aggression: Relationships as social context. *Journal of Applied Social Psychology*, *36*, 2492–2508. doi:10.1111/j.0021-9029.2006.00114.x

Robers, S., Zhang, J., Truman, J., & Snyder, T. D. (2012). *Indicators of school crimes and safety: 2011* (No. NCJ 236021). Washington D.C.: U.S. Department of Education.

Roland, E., & Idsøe, T. (2001). Aggression and bullying. *Aggressive Behavior*, *27*, 446–462. doi:10.1002/ab.1029

Roth, P. L. (1994). Missing data: A conceptual review for applied psychologists. *Personnel Psychology*, *47*, 537–560. doi:10.1111/j.1744-6570.1994.tb01736.x

Roth, P. L., Bevier, C. A., Bobko, P., Switzer, F. S., & Tyler, P. (2001). Ethnic group differences in cognitive ability in employment and educational settings: A meta-analysis. *Personnel Psychology*, *54*, 297–330. doi:10.1111/j.1744-6570.2001.tb00094.x

Rudman, L. A., & Fairchild, K. (2004). Reactions to counterstereotypic behavior: The role of backlash in cultural stereotype maintenance. *Journal of Personality and Social Psychology*, *87*, 157–176. doi:10.1037/0022-3514.87.2.157

Rushton, J. P., & Jensen, A. R. (2005). Thirty years of research on race differences in cognitive ability. *Psychology, Public Policy, and Law*, *11*, 235–294. doi:10.1037/1076-8971.11.2.235

Ryan, R. M., & Deci, E. L. (2000). Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary Educational Psychology*, *25*, 54–67. doi:10.1006/ceps.1999.1020

Salmivalli, C. (2010). Bullying and the peer group: A review. *Aggression and Violent Behavior*, *15*, 112–120. doi:10.1016/j.avb.2009.08.007

Sawyer, A. L., Bradshaw, C. P., & O'Brennan, L. M. (2008). Examining ethnic, gender, and developmental differences in the way children report being a victim of "bullying" on self-report measures. *Journal of Adolescent Health*, *43*, 106–114. doi:10.1016/j.jadohealth.2007.12.011

Schwartz, D., Gorman, A. H., Nakamoto, J., & Toblin, R. L. (2005). Victimization in the peer group and children's academic functioning. *Journal of Educational Psychology*, *97*, 425–435. doi:10.1037/0022-0663.97.3.425

Seals, D., & Young, J. (2003). Bullying and victimization: Prevalence and relationship to gender, grade level, ethnicity, self-esteem, and depression. *Adolescence*, *38*, 735–747.

Sells, L. W. (1976, February). The mathematics filter and the education of women and minorities. Paper presented at Annual Meeting of the American Association for the Advancement of Science, Boston, MA. Retrieved from http://eric.ed.gov/?id=ED121663

Simon, S. (2013, August 21). AP classes failing students. *Politico*. Retrieved from http://politi.co/13QyxQU

Smith, E. (2008). Pitfalls and promises: The use of secondary data analysis in educational research. *British Journal of Educational Studies*, *56*, 323–339.

Smith, P. K., & Brain, P. (2000). Bullying in schools: Lessons from two decades of research. *Aggressive Behavior*, *26*, 1–9. doi:10.1002/(SICI)1098-2337(2000)26:1<1::AID-AB1>3.0.CO;2-7

Smith, P. K., Cowie, H., Olafsson, R. F., & Liefooghe, A. P. D. (2002). Definitions of bullying: A comparison of terms used, and age and gender differences, in a fourteen–country international comparison. *Child Development*, *73*, 1119–1133. doi:10.1111/1467-8624.00461

Sobel, M., Gilmartin, J., & Sankar, P. (2016). Class size and confidence levels among female STEM students. *IEEE Technology and Society Magazine*, *35*, 23–26. doi:10.1109/MTS.2016.2518251

Spears, B., Slee, P., Owens, L., & Johnson, B. (2009). Behind the scenes and screens: Insights into the human dimension of covert and cyberbullying. *Journal of Psychology*, *217*, 189–196. doi:10.1027/0044-3409.217.4.189

Spelke, E. S. (2005). Sex differences in intrinsic aptitude for mathematics and science?: A critical review. *American Psychologist*, *60*, 950–958. doi:10.1037/0003-066X.60.9.950

Spriggs, A. L., Iannotti, R. J., Nansel, T. R., & Haynie, D. L. (2007). Adolescent bullying involvement and perceived family, peer and school relations: Commonalities and differences across race/ethnicity. *Journal of Adolescent Health*, *41*, 283–293. doi:10.1016/j.jadohealth.2007.04.009

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

Stuart-Cassel, V., Bell, A., & Springer, J. F. (2011). *Analysis of state bullying laws and policies* (No. ED-CFO-10-A-0031/0001). Washington D.C.: U.S. Department of Education.

Summers, L. H. (2005, January 14). *Remarks at NBER conference on diversifying the science & engineering workforce*. Presented at the NBER Conference, Cambridge, MA. Retrieved from http://www.harvard.edu/president/speeches/summers_2005/nber.php

Sutton, J., Smith, P. K., & Swettenham, J. (1999). Social cognition and bullying: Social inadequacy or skilled manipulation. *British Journal of Developmental Psychology*, *17*, 435–450. doi:10.1348/026151099165384

Sweeting, H., & West, P. (2001). Being different: Correlates of the experience of teasing and bullying at age 11. *Research Papers in Education*, *16*, 225–246. doi:10.1080/02671520110058679

The College Board. (2013). *Program facts: Overview of the AP program*. Retrieved from http://press.collegeboard.org/ap/fact-sheet

Thomas, S. L., & Heck, R. H. (2001). Analysis of large-scale secondary data in higher education research: Potential perils associated with complex sampling designs. *Research in Higher Education*, *42*, 517–540.

Tulloch, M. (1995). Gender differences in bullying experiences and attitudes to social relationships in high school students. *Australian Journal of Education*, *39*, 279–293. doi:10.1177/000494419503900305

Tyler-Wood, T., Ellison, A., Lim, O., & Periathiruvadi, S. (2012). Bringing Up Girls in Science (BUGS): The effectiveness of an afterschool environmental science program for increasing female students' interest in science careers. *Journal of Science Education and Technology*, *21*, 46–55. doi:10.1007/s10956-011-9279-2

United States Census Bureau. (2012). *School Enrollment*. Retrieved from http://www.census.gov/hhes/school/data/cps/historical/index.html

United States Department of Education. (2013). *Science, technology, engineering, and math: Education for global leadership*. Retrieved from http://www.ed.gov/stem

United States Government Accountability Office (2012). *Science, technology,* *engineering, and mathematics education: Strategic planning needed to better manage overlapping programs across multiple agencies* (Report #GAO-12-108). Washington, D.C.: Author.

Vervoort, M. H. M., Scholte, R. H. J., & Overbeek, G. (2010). Bullying and victimization among adolescents: The role of ethnicity and ethnic composition of school class. *Journal of Youth and Adolescence*, *39*, 1–11. doi:10.1007/s10964-008-9355-y

Wampold, B. E. (2006). Designing a research study. In F. T. L. Leong & J. T. Austin (Eds.), *The Psychology Research Handbook: A Guide for Graduate Students and Research Assistants* (2nd ed., pp. 93–103). Thousand Oaks, CA: Sage.

Wang, J., Iannotti, R. J., & Nansel, T. R. (2009). School bullying among adolescents in the United States: Physical, verbal, relational, and cyber. *Journal of Adolescent Health*, *45*, 368–375. doi:10.1016/j.jadohealth.2009.03.021

Wienke Totura, C. M., Green, A. E., Karver, M. S., & Gesten, E. L. (2009). Multiple informants in the assessment of psychological, behavioral, and academic correlates of bullying and victimization in middle school. *Journal of Adolescence*, *32*, 193–211. doi:10.1016/j.adolescence.2008.04.005

Williams, L. M., & Peguero, A. A. (2013). The impact of school bullying on racial/ethnic achievement. *Race and Social Problems*, *5*, 296–308. doi:10.1007/s12552-013-9105-y

Wout, D., Danso, H., Jackson, J., & Spencer, S. (2008). The many faces of stereotype threat: Group- and self-threat. *Journal of Experimental Social Psychology*, *44*, 792–799. doi:10.1016/j.jesp.2007.07.005

Xie, H., Swift, D. J., Cairns, B. D., & Cairns, R. B. (2002). Aggressive behaviors in social interaction and developmental adaptation: A narrative analysis of interpersonal conflicts during early adolescence. *Social Development*, *11*, 205–224. doi:10.1111/1467-9507.00195

Young, R., & Sweeting, H. (2004). Adolescent bullying, relationships, psychological well-being, and gender-atypical behavior: A gender diagnosticity approach. *Sex Roles*, *50*, 525–537. doi:10.1023/B:SERS.0000023072.53886.86