HIGH-STAKES ACCOUNTABILITY: EXAMINING STUDENT AND TEACHER ANXIETY
WITHIN LARGE SCALE TESTING

by

Nathaniel Paul von der Embse

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

School Psychology

2012

ABSTRACT

HIGH-STAKES ACCOUNTABILITY: EXAMINING STUDENT AND TEACHER ANXIETY
WITHIN LARGE SCALE TESTING

By

Nathaniel Paul von der Embse

The current study examined student and teacher experiences within a high-stakes testing situation by measuring test anxiety and subsequent test performance. The FRIEDBEN Test Anxiety Scale (FTAS) and the State Trait Anxiety Inventory (STAI) were administered to 1465 students at six high schools. In addition, a total of 118 teachers were surveyed from five high schools to examine the presence of teacher anxiety. Data were analyzed to determine differences in state and trait anxiety one week prior to the Michigan Merit Exam, predictors of test anxiety and test performance, the relationship of student career goals to test anxiety, and teacher and student anxiety with respect to school AYP status. Results indicated that trait anxiety was significantly higher than state anxiety. Previous academic achievement, gender and minority status were significant predictors of test anxiety. Test anxiety was a significant predictor of test performance when controlling for significant demographic variables. College bound students reported higher levels of anxiety than did non-college bound students. Students in schools that made AYP reported higher levels of anxiety than did students in schools that did not make AYP, whereas there was no significant difference among teachers in the same schools. This study provides a more nuanced understanding of the test anxiety phenomenon, specifically manifested in a high-stakes environment, by identifying significant predictors of anxiety and test performance. Implications for assessment, intervention and research with test anxiety in a high-stakes context are discussed.

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Nearly one third of the nation's schools have failed to make Adequate Yearly Progress (AYP) as defined by the federal No Child Left Behind law (NCLB, No Child Left Behind Act of 2001); researchers have noted that nearly 80% of schools in most states are not expected to meet these goals by 2014 (Darling-Hammond, 2007). Recent reform efforts (e.g., NCLB) have underscored the importance policymakers have placed on test-based accountability systems. This growing emphasis on accountability attached to large scale tests may be attributed to its relative low cost (as compared to decreasing class size or implementing curricular changes), and to its ability to be externally mandated, implemented rapidly, and provide results which are visible to educators, parents, and policymakers (Linn, 2000). As test-based accountability reforms become more prominent in many educational decisions (e.g., graduation, grade promotion, school effectiveness, ratings, teacher tenure), schools and students are under increasing pressure to meet rising standards for all subgroups culminating in 100% of students meeting proficiency benchmarks by the 2013-2014 school year. Administrators are forced to make tough choices on how to meet these targets. Some schools have resorted to focusing intervention efforts on "bubble" student groups who score close to passing in hope of meeting AYP, as opposed to helping all students across the spectrum of performance (Diamond & Spillane, 2004; Neal, Schanzenbach, & National Bureau of Economic Research, 2007).

Students are routinely tested throughout their academic careers. They are often aware of the stakes tests have on their academic careers and educational decision making (Connor, 2003). Previous research has shown that students are aware and more anxious for high-stakes tests in comparison to typical classroom tests (Segool, 2009). However, it is often assumed that the

1

"stakes" of the test is the same across students and there has been no previous research to examine how test anxiety might vary according to student career goals. For the purposes of this study, "stakes" of a test are defined as the consequences specific to a group of teachers, schools or students (i.e., a test may be considered to have high stakes if a student needs a certain score in order to graduate, if a teacher's evaluation for job tenure is dependent on student test performance, or if a school does not receive government funding for failing to meet annual performance goals) that is directly triggered by performance on tests within accountability policies (e.g., statewide tests required by each individual state educational agency). Test consequences, in contrast to test stakes, may encompass a broad range of effects (both direct and indirect) and are commonly associated with the implementation of testing policies and practices.

There are differences between high-stakes tests and low or medium-stakes tests identified in the literature. These definitions traditionally have considered the external stakes for groups of people rather than how an individual taking the test may perceive test consequences. A high-stakes test has been defined as any test which involves consequences for students or schools; these consequences may include public reporting of schools not making adequate yearly progress (AYP) goals, denial of a student's high school diploma, or failure to qualify for scholarship money (Cizek & Burg, 2006; Heubert, Hauser, & National Academy of Sciences - National Research Council, 1999). The ACT or SAT are sometimes considered medium-stakes tests because stakes (such as university admission or non-admission depending on student academic goals) are delayed and not used in grade promotion decisions (Cizek & Burg, 2006). Another definition within the research literature describes a high-stakes test as having serious consequences for students (e.g., denial of diploma) , medium-stakes as having delayed, indirect consequences for students, and low-stakes as having no consequences for students, yet still may

have consequences for schools (Braden, 2007).

However, there has been little research on student and teacher responses in these "high-stakes" situations to assist in understanding the perceived importance of the test and its consequences for these individuals. Broad generalizations are made concerning the consequences associated with high-stakes tests; a more nuanced analysis may reveal differences among student and teacher groups related to the perceived stakes of the test. Previous definitions (Braden, 2007; Cizek & Burg, 2006) have focused on assumed external consequences of these tests for students and teachers. Furthermore, it is important to recognize that students may have vastly different perceptions of the consequences of the test. For instance, the ACT may be high-stakes for some college bound students but not as important for a student interested in entering the workforce or attending a community college.

Student test performance has become even more important to a variety of stakeholders, including teachers, with increasingly more states adopting policies requiring the use of student test scores in evaluations for teacher tenure. For example, legislators in Michigan recently approved the use of student test scores in the evaluation for teacher tenure (M. Mason, personal communication, December, 2010). Moreover, newspapers such as the New York Times (Santos & Otterman, 2012) and Los Angeles Times (Song & Felch, 2011) have published the results of student test scores as a way to rank the effectiveness of teachers in the public school system. These recent changes, including the public reporting and ranking of teachers, have arguably increased the pressure on teachers regarding the performance of their students on statewide tests (Briggs & Domingue, 2011). As these practices become more widespread, it bears watching if teacher anxiety and subsequent student anxiety may increase as a result. Moreover, school administrators may be increasing the pressure on teachers to raise student test scores knowing

3

that performance data will be made public knowledge in a variety of outlets (e.g., newspapers, television stations).

Test anxiety has risen with the increased use of high-stakes testing in educational decision making (McDonald, 2001; Putwain, 2007, 2008b). Test anxiety is defined as the physiological, psychological and emotional responses to a threatening evaluative situation or stimulus (Zeidner, 1998). Researchers have estimated that 30% of all students suffer from various levels of test anxiety (Gregor, 2005). Students with high levels of test anxiety have been shown to score lower on examinations and have lower grade point averages (Cizek & Burg, 2006; McDonald, 2001; Sena, Lowe, & Lee, 2007). Debilitating anxiety is negatively correlated with test performance (Rafferty, Smith & Ptacek, 1997). In addition, students with disabilities, women, and minority students have higher rates of test anxiety (Putwain, 2007, 2008c; Rosairio, et al., 2008; Sena, et al., 2007; Zeidner, 1990). However, previous research has not considered the direct consequences of the test for individual students. This study examined student and teacher perceptions of the testing situation to identify the *individual* stakes of the test. If test anxiety is differentially correlated with test performance by various demographic predictors (e.g., socio-economic status, gender, special education status, minority status), we may begin to have a more complete understanding of which student subgroups may be the most susceptible to test anxiety and test underperformance. This understanding could lead to targeted intervention services intended to minimize the negative relationship between test anxiety and test performance, ultimately improving performance for students, teachers and schools.

Current Study

The use of high-stakes tests in educational accountability systems can have a range of consequences for students, teachers, schools and the community at large. However, there are

several variables which may relate to the test experiences of students including student test anxiety, career goals, school AYP status, teacher anxiety and demographic variables. Teacher anxiety may play a role in the manifestation of student anxiety as teachers may inadvertently communicate their nervousness to students (e.g., a teacher may indicate to the students that "it is very important to the school and myself that you all perform well on this test"). Test anxiety may be a disruptive force preventing certain student subgroups from performing to their true capabilities as measured on a high-stakes test.

Given the high-stakes associated with these examinations for schools, teachers, and students, it is imperative to understand the relationship of these variables and their potential influence on test performance to ensure the authentic measure of student achievement and school effectiveness. Additionally, this research could inform future measurement and accountability policy as well as help to target future intervention in a meaningful manner. Results may necessitate a careful analysis of the weight placed in test scores in all types of educational decision making, especially if there is initial evidence for potential biases or contributions to achievement gaps. Finally, a closer examination of what constitutes high-stakes tests for students and teachers is needed to inform future research and policy. There is a void in the research literature that examines the relationship of these variables with performance on a high-stakes test at the high school level. The goal of this research is to begin to fill this void.

Research needs to be conducted to deepen our understanding of the student experience within a high-stakes testing environment. This research is important due to the enormous weight placed on student test scores in many educational decisions and the need to identify any variables (e.g., test anxiety) which may interfere with the authentic assessment of student achievement. The study examined student responses on test anxiety scales and subsequent scores on the

Michigan Merit Exam with respect to demographic predictors, student educational goals and

perceptions of test outcomes, teacher anxiety, and school AYP standing. This study explored

how students and teachers perceive the testing experience on the Michigan Merit Exam (MME).

Results from this study expanded the literature base on test anxiety, identifying the individual

"stakes" of a test and the effects of test-based accountability for students and teachers.

CHAPTER 2

LITERATURE REVIEW

Standardized, large-scale tests are frequently being used as a measure of student achievement and the educational quality of a school. The increased emphasis on the data derived from these large scale tests to make important decisions, such as meeting Adequate Yearly Progress (AYP), necessitates a more complete understanding of student and teacher experiences in this testing context in addition to an examination of variables which may influence test performance. Some variables, such as test anxiety, are associated with lower performance on high-stakes tests (Putwain, 2008). This literature review provides a context for this study by highlighting the consequences associated with high-stakes accountability policy for schools, students and teachers, defining test anxiety, and examining and critiquing previous test anxiety research. The rationale for the current study will be s presented in addition to research questions and hypotheses.

Historical Background of High-Stakes Educational Measurement

Since the early 1900's, educators have been looking for ways to measure progress and learning in schools. Early in this era, large scale tests were often used as high school entrance exams and measures of student performance (Resnick, 1980). These tests were designed to provide information on student learning and school effectiveness across multiple grade levels (Koretz & Hamilton, 2006). Some of these tests, such as the Iowa Test of Basic Skills, were intended to help teachers adapt their teaching methods to meet the needs of their students.

In 1965, the Elementary and Secondary Education Act (Title I) was passed to create accountability for schools. This law mandated that schools report student educational progress to receive federal money. In 1965, the National Assessment of Educational Progress (NAEP) was

created and then first implemented in 1969 to monitor the performance of schools nationwide. There was also a marked shift in the purpose of testing in the 1970's, to minimum competency assessments. The minimum competency testing movement sought to ensure a basic level of competency for all students; this also marked the beginning of measurement driven instruction, which was shaped by the belief that outcomes of instruction should meet set standards measurable by a test (Resnick, 1980). The use of large scale tests led some policymakers and the public to view the associated scores as the primary indicators of school success (Koretz & Hamilton, 2006). The NAEP and Title I legislation were national efforts that ushered in the first use of tests as academic monitoring devices and were the precursors to the current use of tests as tools of educational accountability (Hamilton, 2003).

The use of tests to measure student and school progress continued throughout the 1970's and 1980's. During this time consequences were enacted for students who did not pass these exams, such as the denial of a high school diploma or failure to advance to the next grade (Resnick, 1980). This was done, in part, to motivate students to do well on assessments with high stakes for schools. These policies were implemented on a state by state basis and were unequal in implementation and influence on student performance (Hanushek & Raymond, 2004). The 1983 report, *A Nation at Risk*, galvanized public opinion on the perceived failings of American schools. This report led to the expansion of the United States Department of Education, in funding, personnel, and power, in addition to a more critical examination of how assessments were used to determine achievement. *A Nation at Risk* also led to increased stakes of large scale tests, including state departments of education administering financial rewards or sanctions to schools based on test scores (Koretz & Hamilton, 2006).

A Nation at Risk outlined rigorous standards and evaluation of student outcomes, thus

creating an impetus for the increased use of high-stakes testing; meanwhile, there was a growing push to compare the achievement of U.S. students with that of international samples through studies such as the Trends in International Math and Science Study, TIMSS (Mullis, et al., 1998). Results from the TIMSS were used to emphasize higher standards and expectations for secondary students (Ediger, 2001). Reports such as that associated with the TIMSS study helped to usher in our modern system of educational accountability (Hamilton, 2003). In 1994, the Improving America's Schools Act (IASA) required each state to establish high performance and content standards and to implement assessment programs to measure performance towards those standards. This act paved the way for the subsequent passage of NCLB and the associated measurement of adequate yearly progress (AYP). As currently conceptualized within the literature, test based accountability systems "involve four major elements: goals, expressed in the form of standards; measures of performance; targets for performance; and *consequences* attached to schools' success or failure at meeting the targets" (Hamilton, 2003, p.28).

The field of education has now entered an age where high-stakes testing has taken a quite prominent role in the lives of children. Test results have been connected to grade promotion, certification, graduation, and denial/approval of services. Many students are assessed multiple times with government mandated tests throughout their academic careers, often having to pass an exam before advancing in their academic program (Gregor, 2005). The standards-based reform movement gave new, increased emphasis on test-based accountability for schools and students when President George W. Bush signed into law the No Child Left Behind Act (NCLB, No Child Left Behind Act of 2001). NCLB required demonstrable (i.e., measurable) academic progress by states and school districts for all students. NCLB has required rigorous assessment of student outcomes, increasing the importance placed upon student test performance. NCLB

has also resulted in a noticeable shift in control away from the local community to state and federal levels (Hursh, 2006).

High-stakes tests have now become the main indicators of student progress and school effectiveness since the passage of NCLB (Koretz & Hamilton, 2006). Schools across the country are required to meet adequate yearly progress indicators, culminating in 100% of students reaching the set targets by the 2013-2014 school year. If these schools do not meet the required yearly performance targets, they often face a wide range of government mandated sanctions. These accountability policies have changed how schools view student test performance data; this has led to schools increasing the focus of curriculum and intervention efforts on raising student performance on high-stakes tests (Hursh, 2006). While test-based accountability reforms have led to some achievement gains, these improvements were largely the result of student effort and test specific skills at the expense of increased special education placements, preemptive retention decisions and less time devoted to social studies and science (Jacob, 2005).

A school's progress towards AYP is determined by student performance on the science, mathematics, and language arts sections of the test, in addition to the percentage of students taking the test. High-stakes consequences are directly tied to test performance for schools receiving Title I money. Sanctions increase for each successive year of not meeting AYP; this includes not meeting AYP for different subgroups on different sections of the test. Examples of subgroups measured under NCLB requirements include students with disabilities, English Language Learners, students in poverty and racial-ethnic groups. After two consecutive years of not meeting AYP, schools are required to inform parents. After three years, schools must offer supplementary educational services. After four plus years of failure, schools are required to plan systemic action such as whole staff replacement and school restructuring, while such plans must

be implemented after five years of not meeting AYP. The result of not meeting AYP targets can thus have serious consequences for students, teachers, and schools.

Despite questions regarding the reliability and validity of evaluating teachers based upon student test scores (Rothstein, 2010), recent federal initiatives such as Race to the Top (RTTT), have encouraged state departments of education to use student test scores as a way of evaluating teacher tenure applications. RTTT (2009) was the largest competitive grant program in the history of education in the United States (Nicholson-Crotty & Staley, 2012). State departments of education submitted applications that detail "innovative" ways to use student data, with a strong emphasis on using test scores to evaluate teachers. Critics have charged that RTTT is a flawed program which centralizes education authority at the federal level, thus limiting the flexibility and choice of states and local school districts (Onosco, 2011). Others have stated that the RTTT program is unscientific and not research based (Mathis, 2011). Proponents have lauded the program's focus on teacher quality and improving outcomes for all students (Walsh & Jacobs, 2009). One of the main benefits of a winning RTTT application, in addition to a large amount of grant money, was the ability for state departments of education to opt out of AYP requirements and subsequent sanctions.

An Ecological Context for Accountability and Test Anxiety

Bronfenbrenner (1979) developed the ecological systems theory which provides a context for the relationships posited in this study among policy, schools, teachers and student test anxiety. This theory proposes that individual development is affected by nested environmental systems with bidirectional influences both in and between systems; Bronfenbrenner's work is the basis upon which the conceptual framework (See Figure 1) and hypothesized relationships are drawn. This theory focuses on the context of an individual's development within evolving

11

systems and multiple levels of influence. Bronfenbrenner proposed five environmental systems from microsystem (system closest to the individual) into mesosystem (connects with microsystem), exosystem (defines larger social system), macrosystem (outermost layer in individual's development including cultural values, customs and laws), and finally chronosystem (dimension of time). The ecological systems theory highlights the importance of interactions among the macrosystem (e.g., test-based accountability policy, or culture in which individual lives),mesosystem (e.g., school, or social setting which may have an indirect effect on individual), and microsystem (e.g., teachers, settings which directly interact with the subject) and how these interactions may facilitate (or hinder) the development of test anxiety. The following review of the consequences (intended or unintended) of accountability is considered within multiple levels of influence.

Accountability Impact on Schools

For schools that do not meet AYP for several years in a row, there are severe penalties, including being taken over by the state department of education and/or staff losing jobs (Ediger, 2001). In all states, school districts are required to offer vouchers to the parents of students attending schools that do not meet AYP or are classified as in a state of "academic emergency." This policy can potentially result in less money and fewer resources for school districts to address issues affecting student achievement and test performance. Even in schools that are high performing, there is increased pressure to perform better than in years previous and to ensure that all student subgroups are meeting targets.

High schools in particular face unique challenges in making AYP. High schools in Michigan are required by law to disaggregate test score data by any subgroup greater or equal to 30 students. While most elementary schools do not have the requisite 30 students per subgroup

required for reporting and measurement in Michigan, larger high schools do and therefore are responsible for the continued progress of more of these respective subgroups (*At a Glance: NCLB and High Schools. Policy Brief*, 2006). Additionally, researchers have found that the strongest predictor of low performing high schools' AYP status is the number of student subgroups for which they are responsible (Balfanz, Legters, West, & Weber, 2007). These challenges warrant a closer examination of student and teacher variables surrounding the high school, high-stakes examination.

Schools are more likely to determine student progress with a single measure (such as a statewide test) based on the stakes attached to that particular test (Clarke, et al., 2003). In this study, Clark and colleagues interviewed nearly 400 teachers from three different states in which each had three different levels of stakes attached to their assessments (Clarke, et al., 2003). Interviews revealed significant differences in how teachers reported using their assessment data. Results suggest that schools with higher stakes tests are more likely to indicate the academic progress of their student body through the use of a single test; schools with lower stakes tests (those which do not involve grade promotion or retention) are more likely to indicate progress through a variety of sources, including teacher feedback and portfolio assessments.

Accountability Impact on Teachers

There have been a large number of research studies detailing the variety of consequences of high-stakes accountability systems on teachers, including the narrowing of curriculum, changing of instructional practices, changes in teacher morale/anxiety, and changes in teacher retention (Hamilton, et al., 2007; Hamilton, 2003). Smith (1991) in a series of qualitative interviews examined the influence of external testing programs on teacher morale and anxiety. Interview data indicated a significant number of teachers reporting negative emotions, including

anxiety, as test scores were made public. Concerns were raised as teachers reported a belief that test scores were used against them. This study predated the implementation of test-based accountability policies such as NCLB.

Researchers have examined the impact of high-stakes accountability on several schools' climate (Pedulla, et al., 2003). In this study, over 4,000 teachers were asked how high-stakes accountability had affected their morale and school climate. An 80 item likert scale was developed which asked how large scale testing has influenced school climate, pressure on teachers, alignment of classroom practices with state tests, impact on content and mode of instruction and test preparation.

Results generally indicated that the higher the stakes of the test, the greater influence the test had on a variety of areas. Teachers in high-stakes states were significantly more likely than teachers in medium to low stakes states to report student pressure about the test. Teachers in high-stakes states reported significantly more teacher anxiety than teachers in low or medium stakes states. High levels of teacher anxiety were reported regardless of high-stakes for the school or high-stakes for the student. All teachers reported similar, high levels of pressure from parents to raise student test scores. In addition, elementary school teachers were more likely than high school teachers to report instances of student stress in response to the high-stakes assessment. However, in both cases, increased student anxiety did not seem to have a negative influence on teachers' perceptions of school climate (Pedulla, et al., 2003). The identification of stakes within the sample was assumed based on state accountability policy and did not include school specific sanctions (i.e., did the teachers at school XYZ face specific sanctions or did their school consistently meet AYP targets). A closer examination is needed to determine the relationship of *individual* school AYP status (rather than *external* state specific stakes) on

14

teacher reported anxiety. AYP status may also play a role in the expression of teacher anxiety. For example, teachers may be more anxious in a school that is one year (or test cycle) away from receiving severe sanctions. Likewise, teachers may be less anxious in a school that has consistently made AYP, or they may be more anxious given external pressure to continue to make AYP.

Teacher anxiety may also play a role in the manifestation of student test anxiety. Doyle and Forsyth (1973) examined test anxiety levels with a sample of 234 elementary students and 10 teachers. Results indicated a significant relationship between teacher and student anxiety levels. However, the small, non-representative sample raises questions about the validity of conclusions. Another study, with a sample of 1000 Australian students and 32 teachers resulted in no significant relationships among teacher and student anxiety levels (Stanton, 1974). Both studies used relatively simple statistical methodology (e.g., correlations) which may have not fully explained the relationships or lack thereof. Moreover, the two studies were conducted nearly 40 years ago and much has changed in the theory and measurement of student test anxiety. Additionally, the implementation of test-based accountability policies (e.g, AYP and NCLB) creates a new context for the study of the relationship of teacher and student anxiety. Given the sanctions associated with not making AYP, teachers and students arguably have much more pressure to perform on large scale tests than before the enactment of NCLB. Therefore, it is important to examine any contributing factor to the manifestation of anxiety, in both students and teachers that may influence test performance.

Accountability Impact on Students

High-stakes accountability practices have several documented relationships with student variables, from increased achievement (Hanushek & Raymond, 2005) to higher dropout rates

(Amrein & Berliner, 2003). As indicated within the Bronfenbrenner ecological model, accountability policy may indirectly (or directly) influence student behavior. Consequences from high-stakes tests may be magnified for students with disabilities. While raising standards and expectations for students with disabilities is intended to result in positive consequences, some researchers have argued that the implementation of high-stakes accountability has led to an increase in overall special education referrals (Fielding, 2004). School administrators could feel pressure to have students with disabilities not included in the general assessment due to fears of lowered scores and increased sanctions; however, research has shown that schools actually receive increased fiscal and service support to help include these students (Ysseldyke, et al., 2004). Other researchers have found that high-stakes accountability has led to an increase in the scores of students with disabilities and the participation of students with disabilities in assessments, but also to an increase in student anxiety (Katsiyannis, Zhang, Ryan, & Jones, 2007). As previously noted, accountability policies have had both direct and indirect influence on a host of student variables. The proposed conceptual framework (Figure 1), based on the Bronfenbrenner ecological framework, organizes these relationships from government accountability policy (as indicated through school AYP status) to teacher and student anxiety levels. The review of test anxiety to follow describes the history of test anxiety research, current models, and future directions.
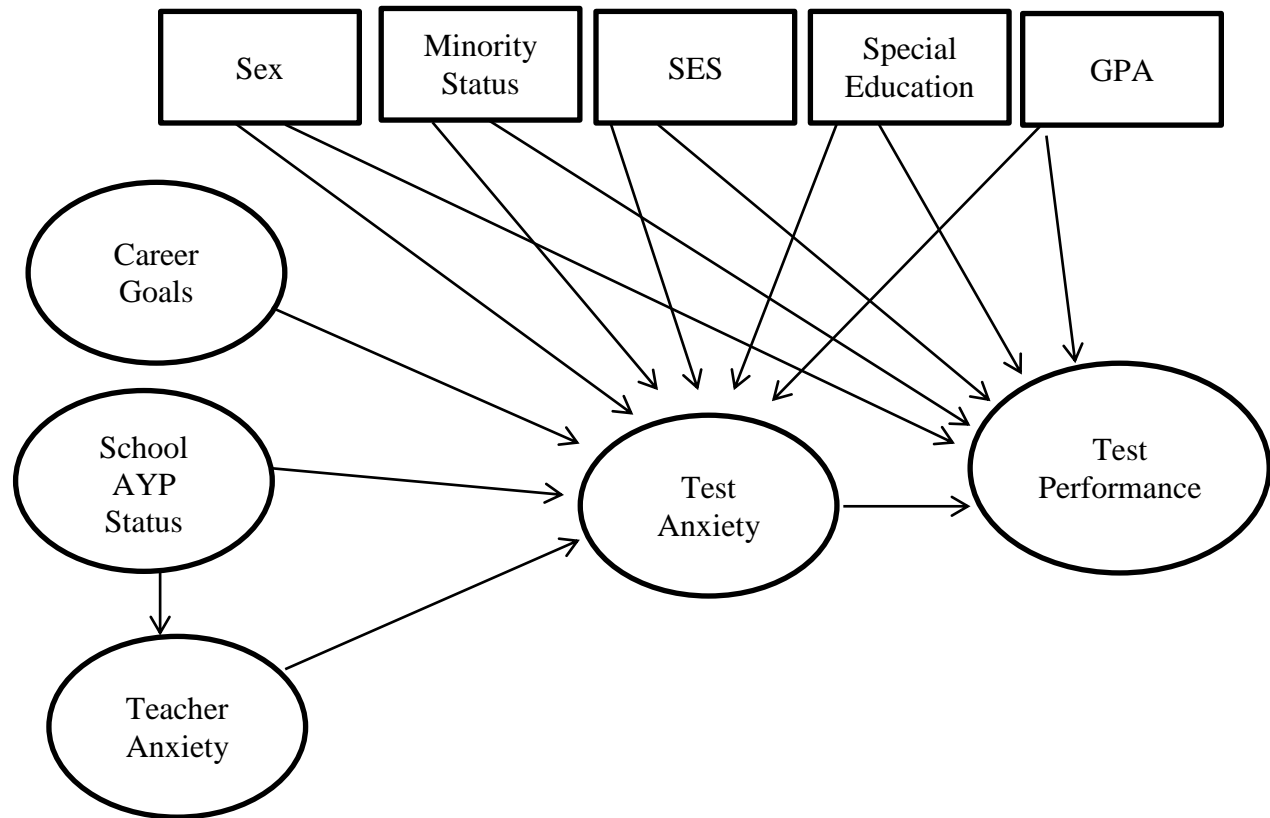
Figure 1: Conceptual Framework

In 1908, Yerkes and Dodson introduced the idea that there is an optimal level of arousal for performance (Yerkes & Dodson, 1908). The researchers suggest that moderate levels of arousal resulted in higher performance, whereas high and low levels of arousal led to decreased performance (i.e., a curvilinear relationship between test anxiety and test performance for individual students). The idea of a tipping point between arousal and performance would influence the next century of test anxiety research. This seminal work predates the study of test anxiety and continues to serve as a theoretical model for many studies.

The first studies which investigated the concept of test anxiety took place as early as 1914 (Folin, Denis, & Smillie, 1914; Stöber & Pekrun, 2004). However, it was not until the 1950's that "test anxiety" was defined and researchers began to investigate the relationship between test anxiety and performance. Sarason and Mandler were pioneers in early test anxiety research as they published a series of studies which demonstrated a negative, *linear* relationship of test anxiety with test performance (Mandler & Sarason, 1952; S. B. Sarason & Mandler, 1952). The Test Anxiety Scale for Children (TASC) was developed by Sarason and colleagues and quickly became the "gold standard" to which all future test anxiety scales would be measured (Seymour B. Sarason, Davidson, Lighthall, & Waite, 1958). A study by Sarason and colleagues investigated test anxiety and test performance for a cohort of children across several years; low test performance was associated with test anxiety and gave additional evidence to the validity and reliability of the TASC (S. B. Sarason, Hill, & Zimbardo, 1964).

The 1960's and 1970's ushered in several important advances in the study of test anxiety. First, a distinction was made between state and trait anxiety in the development of the State Trait Anxiety Inventory (STAI) (Spielberger, 1966). Trait anxiety refers to a prevalent sense of

anxiety that is a continuous syndrome within an individual; state anxiety refers to a temporary

status or symptoms arising from a specific situation or stimulus (Legrand, McGue, & Iacono,

1999). Under normal, non-high stakes conditions, there should be no difference between state

and trait anxiety (Joesting, 1975). Legrand and colleagues (1999) examined heritability and

environmental factors of state and trait anxiety within 547 pairs of twins. Nearly 45% of the

variability in trait anxiety could be accounted for by genetic factors from comparing scores of

monozygotic (MZ) and dizygotic (DZ) twins. Differences on state anxiety scores were non-

significant between MZ and DZ twins; researchers concluded that state anxiety was attributable

to environmental rather than heritable factors. The setting of the study was conducted in an

unfamiliar clinic "thought to provoke anxiety" in the subjects; there was no *high-stakes* test used

as a point of reference for neither item response nor examination within a natural setting such as

a school (Legrand, et al., 1999). Additional research is therefore needed to identify state anxiety

prior to a high-stakes test in a school setting. The STAI is an anxiety instrument which separates

temporary state anxiety (i.e., that which may be associated with an upcoming test) from

continuous trait anxiety. Spielberger would later go on to design several important test anxiety

scales which are commonly used in research today (Spielberger & Vagg, 1984).

The second advance in test anxiety research offered a distinction between the two basic

experiences of test anxiety, worry and emotionality (Liebert & Morris, 1967; Morris & Liebert,

1970). Worry involves the cognitive processes associated with test anxiety, whereas

emotionality includes physiological symptoms. Worry was found to be more related to low test

performance than emotionality, suggesting that cognitive factors have a greater negative effect

than physiological symptoms. These studies continued to influence test anxiety research and test

anxiety scale development for the next 30 years. Test anxiety research would peak in the 1980's

and then substantially decrease in number of publications, which continues today (Stöber & Pekrun, 2004; Zeidner, 1998).

In the late 1990's, Isaac Friedman and Orit Bendas-Jacob developed a new model for measuring test anxiety by introducing a social component (Friedman & Bendas-Jacob, 1997). The FRIEDBEN Test Anxiety Scale (FTAS) was normed on high school and junior high students and moved beyond cognitive and physiological aspects of test anxiety by including the concept of social derogation. Social derogation is a component of test anxiety which identifies the influence of external pressures (i.e. parental or peer pressure) on the rate of test anxiety. The social aspect of test anxiety would prove to be important, as other researchers have noted the impact of peer reference groups and outside pressures on the rates and severity of test anxiety (Goetz, Preckel, Zeidner, & Schleyer, 2008; Marsh, Trautwein, Ludtke, & Koller, 2008). The FTAS has been used in current test anxiety research (Peleg, 2009) and is an important measure now that students are faced with increased pressures from high-stakes accountability systems (McDonald, 2001). However, like most widely used test anxiety scales (e.g. Children's Test Anxiety Scale, Test Anxiety Inventory), the FTAS provides a general measure of test anxiety and not anxiety specific to an upcoming test. Future research could differentiate temporary state anxiety prior to an upcoming high-stake test in addition to providing a more global test anxiety measure.

Research has suggested that increased test anxiety several days before an examination (Raffety, Smith, & Ptacek, 1997). Rafferty and colleagues (1997) examined anxiety prior to a college examination, during, and immediately after the examination. The authors created a scale, the Definitional Anxiety Inventory (DAI), to measure test anxiety ten times over the course of a two week period. Participants recorded responses in an anxiety diary. Profile analysis of

20

repeated measures evaluated anxiety trends, including flatness, level, and parallelism. Debilitating anxiety negatively correlated with examination score, $R(157) = -.25$, $p < .01$. Additionally, levels of anxiety changed over time, $F(9,149) > 18.45$, $p > .001$. Anxiety peeked two days prior to the examination, and dropped during and after the examination. However, the DAI did not specifically indicate the upcoming test as the potential stressor within the protocol items. Further, the DAI did not incorporate social measures of test anxiety, and the sample included only 4th year undergraduate students. The selective nature of the sample (i.e., only included college students), and no indication of student perceptions of test consequences limit generalizability. Additional research is needed to identify levels of *state* anxiety associated with a high-stakes test at the high school level.

Individuals may experience different levels of test anxiety *during* the testing experience (Schutz & Davis, 2000). Researchers asked students through self-reflective questionnaires about their experiences during a test. Results indicated that students must identify the relevance of test stakes for test anxiety to emerge (Schutz & Davis, 2000). Results indicated higher anxiety in relation to test items which were unknown to the student. The measures used in this study were qualitative and did not indicate construct validity with any established test anxiety scales. This has led researchers to develop scales which measure the regulation of arousal or emotions during the testing experience (Schutz, Distefano, Benson, & Davis, 2004). Additional research is needed to quantitatively identify levels of test anxiety corresponding to the importance of test stakes.

Triplett and Barksdale (2005) examined how students perceive high-stakes testing situations. Student responses were in writing and drawing formats; the authors coded responses into one of nine separate categories. Results from surveys of 225 students in several states

indicated that students have an overwhelmingly negative attitude towards high-stakes

assessments, which resulted in increased levels of anxiety and nervousness (Triplett &

Barksdale, 2005). Students reported feeling a sense of isolation during testing situations. The

authors did not have a predetermined coding scheme; rather, they grouped responses by themes.

Inter-rater reliability was not reported for their coding scheme. The drawing and writing method

could be a new way for educators to assess young children's perceptions of high-stakes testing

situations; however, until reliability and validity can be established, results should be used in an

exploratory rather than diagnostic manner.

Despite the overall decreases in test anxiety research (Zeidner, 1998), there continues to

be important advances which expand our traditional understanding of test anxiety. Research on

test anxiety has now examined biopsychosocial factors and social components that influence test

anxiety and test performance (Sena, et al., 2007). The biopsychosocial model posits that

biological (e.g., physiological tenseness, level of arousal), psychological (e.g., emotion or

cognitive factors), and social (e.g., parent pressure) factors all contribute to test anxiety. A

model introduced by Lowe and colleagues highlighted behavior, cognition, and physiological

reactions which correlate with test anxiety and subsequent test performance. Their scale, the

Test Anxiety Inventory for Children and Adolescents (TAICA), further advances the field of test

anxiety by including biopsychosocial factors (Lowe & Lee, 2008; Lowe, et al., 2008).

Test anxiety is now defined as the set of psychological, physiological, and behavioral

responses that accompany concern about possible negative consequences or failure on an exam

(Zeidner, 1998). Test anxious students have high anxiety in evaluative situations and view test

situations as personally threatening. Test anxiety has been found in children as young as seven

years old (Connor, 2003; Putwain, 2008b). In addition, test anxiety has increased since the

implementation of high-stakes accountability policies in countries such as England (Putwain, 2007, 2008a, 2008b). England has similar accountability sanctions for schools that do not meet yearly performance targets to those outlined in NCLB. However, results may not generalize to a broader population due to differences in test administration which may invoke different levels of anxiety than tests in the United States (U.S.) For example, high-stakes tests in England are often not tied to high school graduation or grade promotion whereas more than 20 U.S. states (e.g., Ohio, North Carolina, Louisiana) have graduation or exit high school exams. More research is needed to determine rates of test anxiety since the implementation of high-stakes accountability policy in the U.S.

Prevalence and Demographic Patterns

It has been estimated that upwards of 18-20% of children have an anxiety disorder; there is a lifetime prevalence of nearly 30% of all people suffering from anxiety disorders which has been estimated to cost over $42 billion in treatment (Greenberg, Sisitsky, & Kessler, 1999; Kessler, Chiu, Demler, Merikangas, & Walters, 2005). Test anxiety is a part of overall anxiety and prevalence estimates have varied from as low as 10% to as high as 40% (Cizek & Burg, 2006; King & Ollendick, 1989; Putwain, 2007; Turner, Beidel, Hughes, & Turner, 1993). However, there is great variation in test anxiety manifestation relative to demographic variables. Recently, Putwain (2007) surveyed nearly 1400 10[th] and 11[th] grade students in Great Britain with the Test Anxiety Inventory (TAI). A factor analysis was conducted to determine the acceptability of the TAI with an English sample. African American students reported significantly higher test anxiety scores than Caucasian students ($p < .01$). Socio-economic class, gender, and ethnicity were all strong predictors of test anxiety (both total and components of TAI), but accounted for a relatively small amount of the variance ($F_{11, 1303} = 14.67$, $p < .01$; $R^2 = .09$). The study could

23

have been improved by including previous academic achievement, which is a major determinant

of test anxiety (King & Ollendick, 1989). Findings were consistent with other research

indicating socio-economic status as salient variables in the expression of test anxiety (Hembree,

1988). However, other studies have found no ethnic group test anxiety differences (Zeidner,

1990) and have cautioned against making cross cultural comparisons. Differences in the

classification of "ethnicity" and "socio-economic status" among cultures make comparisons

difficult.

The prevalence of test anxiety in minority students was investigated by Biedel and

colleagues (Beidel, Turner, & Trager, 1994; Beidel & Turner, 1988; Turner, et al., 1993).

Clinical correlates of test anxiety were compared among African American and Caucasian school

children. Over 50% of African Americans (N=27) in the sample reported clinical levels of test

anxiety compared with 38% (N=54) of white students; the authors noted the potential for

overestimation of clinical levels of test anxiety and referred to studies using larger data samples

than their own (N=195). Results indicated no significant differences on test anxiety scores based

on race, $(F_1 = 2.59$, p > .05) (Beidel, et al., 1994). Caucasian students were recruited from a

predominantly (95%) white school, whereas African American students were recruited from a

predominantly African American (90%) school district. Subjects were selected by screening

with the Test Anxiety Scale for Children; the authors did not indicate where the screening took

place, which may have influenced test anxiety scores (Goetz, et al., 2008). It is clear that

additional research needs to be conducted to determine the relationship of demographic variables

such as ethnicity to test anxiety.

Higher rates of test anxiety were found in females and minorities in a meta-analysis of

over 500 test anxiety studies (Hembree, 1988). Hembree calculated mean effect sizes for

differences in test anxiety between males and females; differences were moderate (μ=0.43) in fifth through tenth grades and small (μ =0.27) in eleventh and twelfth grade (Hembree, 1988). Hispanic students reported greater rates of test anxiety than Caucasian students (μ =0.36), whereas differences among African American students and Caucasian students were small (μ =0.21) in early grades and non-existent in high school. This lends support to the findings of higher rates of test anxiety among minorities reported by Turner and colleagues (1993). However, this meta-analysis was conducted prior to the implementation of the high-stakes accountability movement (i.e. No Child Left Behind); Putwain (2008) has suggested that test anxiety has increased since this implementation. Test anxiety in high-stakes testing situations may disproportionately affect special education students. In a recent study, researchers examined the relationship between students with learning disabilities (LD) and test anxiety (Sena, et al., 2007). The study consisted of nearly 800 students with and without specific learning disabilities. The authors concluded, after examination of the factor structure, that the Test Anxiety Inventory for Children and Adolescents (TAICA) was an acceptable instrument for assessing test anxiety in students. Multiple regression analysis was used with LD, gender, and age as predictor variables for the overall test anxiety and multiple subscales of the TAICA in separate analyses. Results of these analyses indicated LD predicted higher cognitive obstruction and inattention on scales of test anxiety. Students with learning disabilities demonstrated higher levels of test anxiety on specific subscales, but not total test anxiety, than their non-disabled peers. However, the predictor variables accounted for only 1-5% of the variance in the criterion variables. Due to the small, relatively homogeneous sample based on convenience, replications are needed across other disabilities to determine test anxiety prevalence and relationship to

achievement. Additional research would strengthen suggestions that educators should assess and potentially accommodate for test anxiety in students with learning disabilities.

Test anxiety also has a social component. The "big fish, little pond" effect demonstrates that children have different levels of anxiety based on the perceived achievement of their peer reference group (Marsh, 1987; Marsh, et al., 2008). Gifted students were found to have higher rates of test anxiety with a gifted peer reference group than gifted students with a non-gifted reference group (Goetz, et al., 2008). When controlled for individual achievement, test anxiety increased with the ability level of the peer reference group. In a sample of nearly 800 Israeli gifted students who were placed in gifted classrooms, individual achievement was significantly negatively (β= -0.16) related to test anxiety; class achievement was significantly positively (β= 0.13) related to test anxiety. Achievement was indicated by teacher-assigned classroom grades. However, it is not clear how "gifted" was defined. Results could be strengthened by incorporating student performance on standardized tests. Further, there is no research which has investigated potential *negative* consequences of social influences beyond peers.

Interviews have shed light on the relationship between career goals and test anxiety. Students' achievement-related motivational beliefs (e.g., career aspirations) were found to be strongly related to higher levels of test anxiety (Ryan, Ryan, Arbuthnot, & Samuels, 2007). However, data were compiled from semi-structured, qualitative interviews with a relatively small sample (N=33). Further, no standardized measure of test anxiety was used. There is a need for a more critical, methodologically sound examination of the relationship between the perceived importance of test consequences and test anxiety.

It should be noted that the majority of studies investigating the prevalence and demographic patterns of test anxiety occurred in non-high-stakes situations. There are few

current studies investigating the prevalence of test anxiety and its relationship to a high-stakes test (Putwain, 2008) following the implementation of NCLB and high-stakes accountability policies. Segool (2009) was one of the first to examine the relationship of test anxiety and high-stakes tests since the implementation of NCLB. Segool explored the differences of test anxiety with elementary school children in response to classroom tests and high-stakes tests. Her sample included nearly 350 students in grades two to five. Results indicated that elementary school children reported significantly higher test anxiety in response to high-stakes tests compared to classroom tests on two different measures of test anxiety. Overall, students also reported higher levels of physiological and cognitive type anxiety on high-stakes tests. Segool also surveyed classroom teachers of the elementary students included in the sample. Teacher reported anxiety was consistent with student reported test anxiety; teachers believed students had significantly higher test anxiety in response to high-stakes tests. Multiple regression analyses revealed that gender and grade were significant predictors of test anxiety whereas ethnicity, general education status and socioeconomic status were not significant predictors. Segool's 2009 study informed the conceptualization of the research questions and analysis in the current study.

Current Study

High-stakes accountability systems, as implemented through NCLB, have a wide range of consequences for schools, teachers and students. Since consequences are determined through the use of high-stakes examinations, it becomes critical to examine any variable which may influence the measurement of outcomes.   As illustrated by the literature review, there is a paucity of research which directly examines the relationship of test anxiety, the "stakes" of a test, and performance on a high-stakes examination. There are methodological weaknesses and contradictory findings which necessitate further study of student and teacher responses to high-

stakes examinations. Segool's (2009) examination test anxiety on high-stakes tests in comparison to classroom tests was a precursor to the current study. That study identified significantly higher test anxiety in response to high-stakes tests and informed the conceptualization of the research questions and methodology used.

Research needs to determine the presence of temporary levels of anxiety which may be associated with an upcoming high-stakes test. Additionally, the relationship between test performance and test anxiety should be carefully examined with respect to demographic factors for a deeper understanding of these relationships within a high-stakes context. This understanding may better inform intervention and services delivered to individuals who may be susceptible to higher levels of test anxiety and therefore perform lower on high-stakes tests. Further, the notion of "high-stakes" needs to be investigated at the student level, with respect to the importance of test stakes. Finally, the relationship of external sanctions to teacher and student anxiety was explored with implications for future accountability and measurement policy. The conceptual framework for this study is presented in Figure 1.

Research Questions and Hypotheses

Question 1. Is there a significant difference between state and trait anxiety one week prior to the MME?

Test anxiety estimates have varied from as low as 10% to as high as 40% (Cizek & Burg, 2006). However, the majority of research demonstrating the prevalence and severity of test anxiety took place prior to the implementation of high-stakes accountability policies such as NCLB. This study sought to identify levels of test anxiety through an examination of students' levels of state anxiety versus trait anxiety. An anxiety instrument (e.g., STAI) was used to differentiate trait type anxiety present on a continuous basis and temporary state type anxiety one

week prior to the MME. Under normal, non high-stakes situations, trait anxiety and state anxiety are expected to be similar as measured on the STAI (Joesting, 1975). Given the high-stakes typically associated with the MME/ACT, students may be experiencing high levels of temporary or state anxiety which may be associated with the upcoming test. The researcher hypothesized that state anxiety would be significantly higher than trait anxiety one week prior to the administration of the MME.

Question 2. Which demographic variables are significant predictors of test anxiety and test performance? What is the relationship between test anxiety and test performance when controlling for various demographic variables?

There have been a number of researchers who demonstrated the relationship between demographic variables and test anxiety and test performance; however, there are few which have examined these variables with high school students in high-stakes situations (Beidel, et al., 1994; Goetz, et al., 2008; Peleg, 2009; Putwain, 2007, 2008c; Sena, et al., 2007; Turner, et al., 1993; Williams, 1996). This study filled this void by examining the relationship of test anxiety and test performance when controlling for various demographic indicators such as ethnicity, socio-economic status, gender, disability status and academic performance. Based on the literature review and previous research identifying significant predictors of test anxiety (albeit in isolation and not considered together within a high-stakes environment), the author hypothesized that women, minorities, individuals of low SES, and students with a disability would indicate higher levels of test anxiety than men, individuals of high SES, Caucasians, and students without disabilities (Hembree, 1988; Putwain, 2007, 2008). In addition, the author hypothesized that test anxiety would be negatively related to test performance even after controlling for various demographic variables.

Question 3. What is the relationship between student career goals and test anxiety?

How students perceive the importance of test consequences may dictate how anxious they are for the examination. As noted in the literature, student test anxiety increases when they are aware of the consequences of the test (Connor, 2003; Schutz & Davis, 2000). Students who plan on going to college may have a different perception of test consequences than students wanting to enter the work force. Students may be anxious on different parts of the test (e.g., students may be less anxious on the math examination which is not directly used for college admission versus the ACT which is a significant determinant in college admission) depending on their career goals. It was hypothesized that students who plan to go to a four year college would indicate higher levels of anxiety on the ACT section than students who are planning on entering the military, the workforce, community college or vocational school. This hypothesis is based upon the belief that the outcomes of the ACT are perceived as more important for the students wishing to attend college and thus more anxiety provoking. This question explores the "stakes" of the test based on student perceptions of the importance of the test and its consequences.

Question 4. What is the relationship between school AYP status and anxiety (teacher and student)? What is the relationship between teacher anxiety and student anxiety?

With the recent Race to the Top initiative, the United States Department of Education is encouraging states to evaluate teacher effectiveness. Student test scores are one of the most popular methods of evaluating teacher effectiveness. In Michigan, the legislature had recently adopted a teacher evaluation system which includes student test scores as a significant portion of the evaluation that determines retention, promotion, and certification (M. Mason, personal communication, December, 2010). The researcher hypothesized that teachers in schools which have not met AYP for four plus years, in addition to rating high levels of importance for student

30

test performance, would report significantly higher anxiety than teachers in schools that have consistently met AYP. Additionally, students in schools that have not met AYP for four plus years would report significantly higher test anxiety than students who are in schools that have met AYP.

Teacher anxiety is another environmental factor which may influence the manifestation of student anxiety. As indicated in the conceptual model (Figure 1) and in the Bronfenbrenner ecological model, there may be multiple variables which may influence test anxiety. Previous studies (Doyle & Forsyth, 1973) have provided initial evidence for the relationship between teacher and student anxiety while others (Stanton, 1974) demonstrated no significant relationships. Given the high-stakes associated with test performance and the possibility of high teacher anxiety in schools that face greater AYP sanctions, it was hypothesized that teacher anxiety would be a significant predictor of higher student test anxiety.

CHAPTER 3

METHODS

Participants

The participants in this study were 11th grade high school students selected from six high

schools, each from different school districts in Michigan. Eleventh grade students were selected

because they are required to take the Michigan Merit Exam. Research findings indicate that test

anxiety levels are stabilized by middle school and high school (Hembree, 1988). High school

students may have the "highest stakes" due to the importance of the ACT portion of the MME

for college admission in comparison to middle or elementary students. The 2010-2011 cohort of

students was the second to take the MME who have had standardized curricular expectations

across the state of Michigan since entering high school (Michigan Department of Education,

2009).

High schools were selected based on their AYP status from the 2009-2010 test scores.

Six high schools were selected from a statewide list such that three schools in the sample each

indicated one of two levels of AYP goal attainment: 0 years of *not* making AYP (i.e.,

consistently have achieved AYP targets) and four years or greater of *not* making AYP. Schools

were selected based on their willingness to be involved in this research and the feasibility of the

researcher to conduct research at the given sites (e.g., no schools in the Upper Peninsula of

Michigan were selected due to time and financial constraints). AYP was considered not met

based on at least two indicators (e.g., percentage of students taking the test, special education

students not meeting math targets). For instance, if a school has not made AYP *only* due to low

percentages of students not taking the test, it was *not* included in the category of four years or

greater of not meeting AYP. The schools which participated in this study in the first category

had made AYP consecutively for a minimum of three years. The schools in the second category had been classified as "Identified for Restructuring—Implementation" which is the highest level of AYP sanctions in Michigan. A total of 1463 students out of a potential 1965 from the six high schools participated in this study (for a total participation rate of 74%). The total response rate is considered excellent (Hopkins & Gullickson, 1992). Three parents indicated to the researcher that they would not like their child to participate in the study, all from the same school. Student participation rates within participating schools ranged from a low of 53% to a high of 88% (See Table 1). All students who were eligible to take the MME were asked to participate in this study. Students who took the alternate assessment were not included. Current law as indicated in NCLB allows for up to 2% of students to take an alternate assessment and be counted as proficient in the state's accountability system (e.g., students with severe cognitive disabilities).

In addition, all high school teachers from each school were asked to complete a brief Internet-based survey concerning teacher anxiety. All teachers were recruited to take the survey and not randomly selected. Teachers were recruited with the following incentives upon completion of the survey: test anxiety reduction intervention resources, a chance to win a $50 gift card, and a school wide report indicating levels of test anxiety and its relationship with test outcomes. A total of 118 teachers completed the Internet-based survey. The total sample size consisted of 70 teachers from schools that made AYP (total teacher population=205) and 48 teachers in schools that did not make AYP (total teacher population=166) for a total teacher response rate of 32%. One school (which did not make AYP) chose not to distribute the survey to the teaching staff.

Table 1

Characteristics of Participating High Schools

School	Setting	AYP Status	Sample	Population	%[a]
A	Rural	Met	240	274	88%
B	Suburban	Met	135	201	67%
C	Rural	Met	272	326	83%
D	Suburban	Identified for Restructuring	400	468	85%
E	Urban	Identified for Restructuring	192	274	70%
F	Urban	Identified for Restructuring	224	422	53%
Total N			1463	1965	74%

[a]%=the percentage of student participants from the same grade, school population.

Measures

FRIEDBEN Test Anxiety Scale

The FRIEDBEN Test Anxiety Scale (FTAS) was used to measure test anxiety. The

FTAS was selected because it advances traditional measurement of test anxiety and reflects

current test anxiety theory by incorporating physiological and social responses to anxiety in

addition to the widely measured worry (i.e., cognitive) and emotionality (i.e., physiological)

components. The FTAS provides a global measure of test anxiety based upon the most current

test anxiety literature (as opposed to several other test anxiety instruments created much earlier;

Cizek & Burg,2006). The FTAS is a 23-item survey designed to measure test anxiety (Friedman

& Bendas-Jacob, 1997; see Appendix A). The three subscales of the FTAS are social

derogation, cognitive obstruction, and tenseness, including a global measure of test anxiety. It is

designed for use among middle school and high school students. Item letters (see Appendix A)

corresponded with the following numbers on the Scantron response form, A-1, B-2, C-3, D-4, E-

5, F-6, whereas 1 indicated "characterizes me perfectly" and 6 indicated "does not characterize

me at all." Several items (9, 10, 11, 12, and 23) were reverse scored to ensure that all responses

were in the same direction. Next, *all* responses (including the aforementioned items) were

reverse coded such that higher scores on the FTAS indicated higher levels of test anxiety.

Examinee responses were summed for a total test anxiety score.

Prior reliability and validity evidence for the FTAS had been obtained using a sample of

nearly 2000 students from both high schools and junior high schools. Internal consistency

estimates of reliability were 0.91 for total scores and .86, .85, and .81 for each of the three

subscales (Cizek & Burg, 2006). The FTAS was validated using current test anxiety measures

such as the Test Anxiety Inventory (TAI) and the Test Anxiety Scale (Friedman & Bendas-

Jacob, 1997). The correlations between total scores on the TAI and the FTAS were .84 and .82 (e.g., for males and females).   .

State-Trait Anxiety Inventory

The State-Trait Anxiety Inventory or STAI (Spielberger, 1970; Appendix D) was used to measure state and trait anxiety. The STAI was selected for this study as it differentiates between temporary or emotional state anxiety versus long standing personality trait anxiety. The STAI has been considered the "gold standard" in identifying temporary levels of anxiety which may be associated with an upcoming stressor (i.e., the high-stakes test). The instrument is written at a sixth grade level and contains twenty questions with four point Likert response options with 1 indicating "not at all," 2 indicating "somewhat," 3 indicating "moderately so," and 4 indicating "very much." (Speilberger, 1970). State and trait anxiety scores are determined by the sum of examinee responses on the first twenty (state) and second twenty (trait) questions. Several items on each scale were reverse coded to ensure all responses were in the same direction (State—57, 58, 61, 64, 66, 67, 71, 72, 75, 76; Trait—77, 79, 82, 83, 86, 89, 90, 92, 95). Standard scores for male and females students for both state and trait anxiety were calculated by comparing each raw score (sum of item responses for state and trait scales) to the corresponding standard score identified in the STAI manual (Spielberger, 1983). Differences have been indicated between temporary state anxiety and long-standing trait anxiety using the STAI (Joesting, 1975). The STAI was found to have strong convergent validity with the Test Anxiety Inventory and the Test Anxiety Scale(Speilberger, 1983). Test-retest reliability was evaluated and found to be 0.4 for state anxiety and 0.86 for trait anxiety (Rule & Traver, 1983).The low reliability of state anxiety was expected as a response to unique situations. When four factors of the scale were extracted (state anxiety present, state anxiety absent, trait anxiety present, trait anxiety absent), internal

consistency was found to be .92, .92, .94, and .95 (Speilberger & Vagg, 1984).

Student Demographic Form

In addition to the STAI and FTAS, a brief demographic survey was administered to students (see Appendix B). This survey identified student demographic variables (e.g., sex, race, parent education/employment, special education status, academic achievement/GPA) in addition to student career goals and post-secondary plans. Grade Point Average (GPA) and special education status were reverse coded for clarity of interpretation (i.e., the *new* codes for GPA indicate higher numbered responses for higher levels of reported GPA and higher numbered responses indicated special education service received). The student survey was administered twice to practice subjects to determine appropriate wording (i.e., eliminating confusing questions, adding/deleting items) and average time of completion. The first test phase occurred with a group of 20 students. Following the pilot administration, wording of items was adjusted and the average completion time was 20 minutes. A second practice administration occurred with four students to ensure readability and comprehension of survey items. Student feedback indicated that the survey was appropriate for the selected sample. Completion times ranged from 14 to 22 minutes.

Additional Student Anxiety Measures

The student test anxiety survey (see Appendix C) asked students how they perceive the MME and ACT. Students were also asked if they were anxious for a particular part of the test (e.g., ACT). Students responded on a 5-point Likert scale ranging from "strongly agree" to "strongly disagree." Item letters (see Appendix C) corresponded with the following numbers A-1, B-2, C-3, D-4, E-5. Item A indicated "strongly agree," item B indicated "somewhat agree," item C indicated "neutral," item D indicated "somewhat disagree," and item E indicated

"strongly disagree." Items 55 and 56 on the student survey were reverse coded (e.g., "I am **not** anxious to take the ACT") to ensure a similar direction as items 45 and 46. Next, items 45, 46, 55, and 56 were reverse coded such that higher scores indicated higher levels of anxiety.

Several measures of anxiety were derived from data on the student test anxiety survey including MME anxiety (items 45 and 55) and ACT anxiety (items 46 and 56). Items 45, 46, 55, and 56 were correlated as estimates of student anxiety with respect to MME and ACT. While significant (MME Anxiety $r=.26$ and ACT Anxiety $r=.25$), correlations for MME anxiety and ACT anxiety were considered weak (Cohen, 1977). Finally, a summary score (e.g., MME anxiety) was calculated from the average of individual responses to the questions (e.g., MME anxiety calculated from items 45 and 55).

Teacher Survey

A teacher survey (see Appendix E) measured anxiety, attitudes towards the use of student performance data and perception of outside pressures (e.g., administrators and parents). The teacher survey was administered online using www.surveymonkey.com. A principal component analysis was conducted to identify a reliable estimate of teacher anxiety and to reduce the multiplicity of observed measures (items 6, 7, 8, 21 and 22 on Appendix E) into a single component. The Barlett test was significant ($p=.000$) indicating interrelationships between variables. This test suggested a redundancy within the data making a principal component analysis worthwhile. The total variance explained table indicated two eigenvalues above one explaining over 70% of the total variance (MME anxiety=46.80% and ACT anxiety=24.90%). Visual inspection of the scree plot confirmed the "elbow" starting with the third component supporting the notion that the first two components significantly account for the majority of variance. Lastly, the component matrix was examined which suggested Component 1 is closely

associated with Q6 (.88), Q7 (.87) and Q8 (.69) while Component 2 is associated with Q21 (.74) and Q22 (.72). Component 1 was an adequate measure of teacher anxiety (C1= .88*Q6 + .87*Q7 + .69*Q8). Because the principal components' distribution is normal, then no assumptions of multivariate normality were violated.

A new component was created to measure total teacher anxiety, C1= .88*Q6 + .87*Q7 + .69*Q8. The total teacher anxiety component (M=7.72, SD=2.36) was examined to determine potential violations of multivariate normality. Neither Shapiro-Wilk or Kolmogorov-Smirnov tests were significant, visual inspection of the histograms revealed only small deviations from normality and the new teacher anxiety component did not have a skew or kurtosis value outside the 2 to -2 range.

School Administrator Survey

A questionnaire (see Appendix J) was administered to test coordinators and school administrators at each of the six schools to examine the context in which students took the test. The questionnaire examined test administration practices among the schools (e.g., "Do your students typically take a test prep course?" "In what setting do the students take the test?"). In addition, participants were asked if test anxiety reduction techniques were taught, and perceptions of staff and student anxiety levels. Questions were open ended and qualitative in nature. Data from this questionnaire provided contextual information regarding the analysis and subsequent interpretation of the following examination of school AYP status, teacher anxiety and student anxiety.

Michigan Merit Exam

The Michigan Merit Exam (MME) is a criterion-referenced test that measures students' skills in core content areas and consists of ACT® plus writing, WorkKeys® Applied

Mathematics, Reading for Information, Locating Information, and Michigan Mathematics, Science and Social Studies tests. The MME is the annual state accountability testing program for high school students that is used to comply in Michigan with the requirements of the No Child Left Behind Act of 2001. The MME is designed to be an accountability measure for school improvement by measuring student achievement and school progress towards AYP targets. The MME is a criterion-referenced test administered to all eleventh grade students in Michigan. Scores on the MME are compiled and reported to the public in the summer following administration of the test.

The test is taken over the course of three school days and measures students' skills in core content areas. Testing typically occurs during the first week of March with retests two weeks later. The MME tests of science, mathematics, and language arts are multiple choice which take approximately 40 minutes each to complete for a total of 2 hours and are comprised of items written to align with Michigan high school standards and items taken from the corresponding ACT subtest area. Total student testing time is approximately 460 minutes over three days. MME scores are reported with the following cut scores required to be labeled proficient (for students in the eleventh grade): Math—1116, Reading—1108, Science—1126, and Social Studies—1129 (Michigan Department of Education, 2011). Reports are provided to the parents of each student with MME scores, performance levels, standard/domain subscores by subject in addition to ACT and WorkKeys scores. The MME Social Studies subtest was used in this study. This subtest was selected as it was the only subtest on the Michigan Merit Exam which did not include ACT content and consisted of questions written for the exclusive purpose of measuring knowledge of social studies based on Michigan content standards (e.g., the MME math subtest include items unique to Michigan content standards *and* math items from the ACT).

ACT

The ACT® is a college entrance exam which is the primary component of the MME. Students taking the MME can have their ACT® scores sent to up to four prospective colleges at no cost as part of an application for admission and/or scholarships. The ACT® consists of five subject tests: English, Mathematics, Reading, Science, and Writing. The ACT® is administered on the first day of testing and takes approximately 3 hours and 25 minutes to complete. Each subtest used a multiple choice response format with the exception of the writing subtest which provides a prompt followed by a space for the examinee to provide a written response. ACT scores are calculated for each of the five subtests and one composite score. Scores range from a low of 1 to a high of 36.

Procedures

Schools were identified based on AYP requirements, demographic variability and their willingness to participate in the research study. The researcher worked with the Michigan Department of Education to obtain a list of public schools to meet one of two AYP criteria for inclusion. This list consisted of approximately 55 school districts. Next, the list was narrowed to include schools with a diverse student body and within a three hour driving distance from East Lansing. The researcher then contacted schools on these two lists; 14 schools declined to participate before six schools were selected. Support from superintendents and principals was solicited prior to recruiting teachers. Test anxiety intervention resources were offered to participating schools after the completion of data collection. A group level summary of student and teacher responses (without identifying information) was provided to school administrators which indicated aggregate student and teacher response patterns, relationship of demographic

characteristics to test anxiety and test performance, suggestions for intervention and comparison of each school with the total included research sample.

Teachers distributed an informational letter (see Appendix G) which was sent home with students to give to parents in January of 2011. The informational letter specifically described procedures which ensured confidentiality of student information, and individual level student test anxiety data was not made available to any school personnel. This letter informed parents that if they would *not* like their child to participate in the research, they should sign the form and return it to school or contact the researcher through email or phone. Out of 1965 students who were eligible to participate in the survey, three parents indicated that they would not like for their child to participate. The reasons given by the parents included a desire for their child to be individually assessed for test anxiety, concerns of reporting anxiety scores to state agency and privacy concerns. The schools were notified and the identified students were not asked to complete the survey.

The FTAS, STAI, and student surveys were administered to students in February of 2011, approximately one week before the MME was taken. The exact date of administration varied between schools ranging from nine days to four days prior to administration of the MME. Teacher surveys were administered via an email link to www.surveymonkey.com one week prior to the MME and continued through March 2011. Prior to the survey administration, students were informed they had the right not to participate in the study. Students were also informed that all individual information collected would be strictly confidential and not be released to parents or school personnel except in summary form. Each student participant received a unique code which included the school name. No students refused to sign the assent form. The cover sheet

(with the student name) was removed from the rest of the survey packet and kept at each school thereby ensuring confidentiality.

School administrators (and the researcher at two of the participating schools) read standardized administration procedures to the students on how to complete the measures (Appendix F). The entire procedure ranged from 20 minutes to 35 minutes. The survey was administered in different settings (i.e., in four schools the survey was administered in a group setting in the school cafeteria/gymnasium, in the other two schools classroom teachers were asked to administer the survey).

The assessment packet was scored via Scantron response form at a separate location (Michigan State University scoring office) and not connected with student names (only the code). No identifying data left the participating schools. The school secretary and administrators matched MME and ACT scores into the database with the student ID number in July of 2011. The database (including the student code and test scores) was then sent to the researcher and matched with test anxiety scores. This method ensured a) student confidentiality, b) the school did not have individual test anxiety scores, and c) the researcher did not have MME scores connected to student names.

Data Analysis

Data were only used for research purposes and no identifying information was collected. All data were stored on a password protected computer and external hard drive. Hard copies of the completed Scantron response forms were destroyed. The researcher examined raw data for errors through visual analysis and descriptive statistics. All research procedures and protocols met the approved Institutional Review Board guidelines at Michigan State University. Table 1 is a summary of statistical procedures, research questions and variables.

Question 1.Is there a significant difference between state and trait anxiety one week prior to the MME?

A paired sample t-test was conducted to examine differences between state and trait anxiety standard scores as reported on the STAI. Anxiety on the STAI was considered a continuous variable.

Question 2. Which demographic variables are significant predictors of test anxiety and test performance? What is the relationship between test anxiety and test performance when controlling for various demographic variables?

A stepwise, multiple regression was performed to examine the relationship among various demographic variables with test anxiety to identify significant predictors (Pallant, 2007). A stepwise, multiple regression was used to examine the relationship among various demographic variables with test performance. A hierarchical regression analysis was used to examine the relationship of test anxiety (total test anxiety on the FTAS) with test performance (MME Social Studies and ACT composite) when significant predictors (as indicated by the prior regression analyses) were included in the model. A power analysis was conducted using the statistical software G*Power 3 and determined that a minimum of N=367 was needed for a moderate effect size (Cohen, 1977). Academic achievement, special education status, minority status, and gender were entered into the first regression block. These variables were selected due to the hypothesized predictive power within previous test anxiety and test performance regression analyses. FTAS total was entered into the second block to examine the unique contribution of test anxiety in predicting test performance.

Question 3. What is the relationship between student career goals and test anxiety?

A multivariate regression analysis was used to examine the relationship between student career goals and test anxiety (total test anxiety on the FTAS, MME anxiety, and ACT anxiety), when controlling for significant predictors as identified in the previous research question. Student career goals were treated as a discrete variable, with student responses grouped into one of two categories: attending a four year college (as indicated on student survey, see Appendix B) or other (including attending a 2 year college, workforce, military or vocational school). This grouping is due to the relative importance of the ACT for students attending a four year college.

Question 4. What is the relationship between school AYP status and anxiety (teacher and student)?

The analyses conducted for question four were exploratory in nature given the small sample size (six high schools). A MANOVA was conducted to determine potential differences in anxiety (MME anxiety, ACT anxiety, and total test anxiety on FTAS) between students in schools grouped on the two different AYP categories. An independent samples t-test was conducted to determine potential differences in teacher reported anxiety (derived from the principal component analysis) in schools grouped by AYP category. A final, exploratory analysis examined the relationship between teacher anxiety and student anxiety. Using the principal component analysis as a measure of teacher anxiety, a mean teacher anxiety score was assigned to each participant matched by school (i.e., students in school A all received the same teacher mean score from school A). This was done because it was not possible to individually match students and teachers nor was it practical (e.g., students in high school often have five or more teachers). Thus, a mean teacher anxiety score was derived and then assigned to the respective participants. A bivariate correlation analysis between student and teacher anxiety was used and if

45

significant, the variables would have been included in a multiple regression equation identified in

Question 2, part 1 to identify significant predictors of student anxiety.

Table 2

Data Analysis

Research Question	Measures	Variables	Data Analysis
1. (a). Is there a significant difference between state and trait anxiety one week prior to the MME?	STAI	STAI-State STAI-Trait	Paired Sample T-test

Research Question	Measures	Variables	Data Analysis
2. (a). Which demographic factors predict test anxiety?	FTAS Student Survey	FTAS Demographic predictors	Stepwise regression
2. (b.) Which demographic factors predict test performance?	MME Student Survey	ACT Composite MME Social Studies Demographic predictors	Stepwise regression
2 (c). What is the relationship between test anxiety and test performance when controlling for significant demographic variables?	MME-FTAS Student Survey	ACT Composite MME Social Studies FTAS Demographic predictors	Hierarchical regression

Table 2 continued.

Research Question	Measures	Variables	Data Analysis
3. (a). What is the relationship between student career goals and test anxiety?	Student Survey FTAS	MME Anxiety (items 45, 55) ACT Anxiety (items 46, 56) FTAS Career Goals	Multivariate regression

Research Question	Measures	Variables	Data Analysis
4. (a). What is the relationship between school AYP status and teacher anxiety?	Teacher Survey	Teacher Anxiety (Principal component) AYP Status	Independent Samples T-test
4. (b). What is the relationship between school AYP status and student anxiety?	FTAS Student Survey	MME Anxiety (items 45, 55) ACT Anxiety (items 46, 56) FTAS AYP Status	MANOVA
4. (c) What is the relationship between student test anxiety and teacher test anxiety?	Teacher Survey FTAS	Teacher Anxiety (Principal component) FTAS	Bivariate Correlation Multiple Regression

CHAPTER 4

RESULTS

Student Sample Demographics

      The sample included 694 male and 738 female students who ranged in age from 15 to 19, with 12.1% students who were currently receiving special education services and 7.8% who were currently receiving gifted education services. The racial composition of the sample was 74% Caucasian, 6.5% Hispanic, 3.3% Asian, 4.7% African American, 1.4% Native American and 8.5% Multiracial. One third of the students in the sample reported receiving free or reduced lunch (32.8%), thus qualifying for "low" SES. Academic performance (as measured by overall Grade Point Average, GPA) was evenly distributed with 24.9% of students indicating GPA between 4.0-3.50, 25.4% between 3.49 and 3.0, 21.3% between 2.99 and 2.50, 14% between 2.49 and 2.0, and 8.8% indicating a GPA between 1.99 and 0. A strong majority of student participants indicated four year college as a career goal (72.7%) compared to other post-secondary plans (e.g., work force, community college, military). Table 3 summarizes the overall demographic characteristics of the students who participated in the study and broken down by school AYP status. In some demographic categories, the total N does not add up to the total 1463 participants which indicated a small proportion of missing responses. Students who did not complete at least 10% of the total survey items or indicated an unusual response pattern (e.g., filled in all C's on the scantron form) were dropped from the sample for a total of 26 students. As indicated in the demographic characteristics on the student response form, these 26 students were not significantly different from the sample population on demographic variables (e.g., sex, ethnicity, age). Specific to responses on the State-Trait Anxiety Inventory, 302 participants were not included in the analysis of state and trait anxiety. These participants were dropped due to

incomplete responses on the STAI (and thereby making the calculation of total state and trait

anxiety challenging). In addition, over 200 participants were not included in the test

performance analysis as one of the sample schools did not provide test score data despite

repeated attempts at contacting the administrator in charge of data preparation and delivery.

Table 3

Demographic Characteristics of Student Participants by AYP Status

Demographic Characteristic		AYP Met N	AYP Not Met N	Overall N	% [ab]
Age	15	6	10	16	1.1
	16	384	442	826	56.5
	17	238	320	558	38.1
	18	16	23	39	2.7
	19	2	2	4	0.3
Gender	Male	305	389	694	47.4
	Female	336	402	738	50.4
Race	Non-minority	523	558	1081	73.9
	Minority	122	235	357	26.1
Free or Reduced Lunch	Yes	166	302	468	32.8
	No	477	484	961	65.7
Special Education Status	Special Education	75	103	178	12.1
	General Education	565	662	1227	83.9
Overall GPA	3.50-4.0	172	192	364	24.9
	3.0-3.49	181	191	372	25.4
	2.50-2.99	149	163	312	21.3
	2.0-2.49	84	111	195	13.3
	1.50-1.99	41	52	93	6.4
	0.0-1.49	16	19	35	2.4
Career Goals	Four Year College	524	539	1063	72.7
	Two Year Community College	83	124	207	14.1
	Vocational School	5	14	19	1.3
	Work Force	5	19	24	1.7
	Military or National Guard	27	46	73	5.0

[a]%=the percentage of student participants from the same grade, school population.

[b]Missing data within each category resulted in total percentages less than 100%

Question 1

Part 1—Is there a significant difference between state and trait anxiety one week prior to the

MME?

It was hypothesized that state anxiety would be significantly greater than trait anxiety one week prior to the MME. Analyses were conducted to determine if there was a significant difference among state and trait anxiety as measured by the STAI. Under normal conditions, state and trait anxiety are expected to be similar (Joesting, 1975). A paired sample t-test was selected as the sample was distributed normally. As expected with a paired sample, the state and trait scores were highly correlated (*r (1155)*=.75, *p*<.01). However, contrary to the proposed hypothesis, the mean trait score (*M*=52.73, *SD*=11.33) was greater than the mean state score (*M*=51.80, *SD*=11.86). A paired-samples t-test showed significance beyond the .001 level: *t* (1155)= -3.54; *p*<.01 (two-tailed). Average raw scores on trait and state anxiety were compared with a comparable percentile rank as indicated in the STAI manual. Results indicated a (raw) trait mean of 41.35 and a (raw) state mean of 39.81 for male students; this corresponded to the 56th percentile for trait anxiety and the 54th percentile for state anxiety. The (raw) state mean of 44.66 and (raw) trait mean of 45.46 for female students corresponded to the 68th percentile and 69th percentile respectively (Spielberger, 1983). Table 4 summarizes the descriptive statistics for the state and trait anxiety scores.

Table 4

State and Trait Anxiety Scores

Gender	Variable	N	*M*	*SD*	Skewness	Kurtosis
Male	Raw State	588	39.82	12.87	.35	-.51
	Raw Trait	571	41.35	12.48	.36	-.15
	Standard State	588	50.34	13.29	.32	-.59
	Standard Trait	571	51.10	11.80	.37	-.15
Female	Raw State	636	44.66	13.10	.15	-.56
	Raw Trait	613	45.46	11.33	.06	-.33
	Standard State	636	53.16	10.19	.15	-.57
	Standard Trait	613	54.25	10.65	.05	-.33
Total	Raw State	1224	42.33	13.21	.23	-.57
	Raw Trait	1184	43.48	12.07	.17	-.31
	Standard State	1224	51.80	11.86	.17	-.49
	Standard Trait	1184	52.73	11.33	.18	-.30

Question 2

Part 1—Which demographic variables predict test anxiety?

The author hypothesized that women, minorities, individuals of low SES, and students with a disability report higher levels of test anxiety than men, individuals of high SES, Caucasians, and students without disabilities. A stepwise regression was used to examine the relationship among sex (Male and Female as indicated by 0 or 1), minority status (Caucasian and non-Caucasian as indicated by 0 or 1), socio-economic status (low indicated by 0 or high indicated by 1 as measured by the receipt of free or reduced lunch), age (from age 15 through 19), academic achievement (Overall Grade Point Average, from 4.0 to 0, with 6 indicating 4.0) and special education status (general education 0, special education 1) with test anxiety (total test anxiety on the FTAS). Several preliminary analyses were conducted for each regression analysis to ensure assumptions were met. No violations of multicollinearity were indicated in variance inflation factor scores (VIFs; VIFs <10) and tolerance (Tols. > .10). A visual analysis of the standardized regression residuals was conducted and found to be normal.

Non-significant predictors were dropped from the overall model using the criteria of $p >$ .10. The original model included sex, minority status, socio-economic status, age, academic achievement, and special education status. Age was dropped from the first model, special education status was dropped from the second model and socio-economic status was dropped in the third model (using the $p > .10$ criteria). The final (i.e., fourth) model included minority status, academic achievement, and sex that were significant predictors and explained 10% (R^2=.10) of the variance of total test anxiety on the FTAS, $F (3, 1309) = 50.38, p < .001$ (see Table 5). Sex had the highest beta level ($\beta = .602, p < .001$) followed by minority status ($\beta = -$.158, $p < .01$) and academic achievement ($\beta = -.09, p < .001$). On average, females scored higher

on anxiety measures than males, students from non-minority backgrounds scored higher on anxiety measures than students from minority backgrounds, and low achievers scored higher on a measure of anxiety than students with high previous academic achievement.

A post hoc analysis was conducted to determine the influence of interaction terms on the significance of the independent variables with total test anxiety. The initial model included all of the 2x2 interaction terms of the independent variables (e.g., age by minority status). A general liner model was created in SPSS which included all of the independent variables plus interaction terms. Results indicated that there were no statistically significant interaction terms after a Bonferroni correction was applied (SES by minority status was initially significant, $\beta = .308$, $p < .05$).

Part 2—Which demographic variables predict test performance?

A second, stepwise regression was used to examine the relationship among the demographic predictors as identified in Part 1 with test performance. Separate stepwise regressions were conducted for two measures of test performance (ACT Composite Score and MME Social Studies) due to the different scales of each dependent measure. The MME Social Studies subtest score was selected as it was the only test score which did not include ACT content and consisted of questions written for the exclusive purpose of measuring knowledge of social studies based on Michigan content standards. Preliminary analyses conducted for each regression analysis indicated no violations of multicollinearity (VIFs; VIFs <10) and tolerance (Tols. > .10). The standardized regression residuals were found to be normal through a visual analysis.

Similar to Part 1, all non-significant predictors were dropped from the model using the criteria of $p > .10$. The hypothesized model was comprised of the following independent

predictors: sex, minority status, socio-economic status, age, academic achievement, and special education status. In the first model (ACT Composite as dependent variable), socio-economic status was dropped resulting in a significant model. The final model explained 39% (R^2=.385) of the variance of test performance on the ACT, F (5, 1050) = 131.66, p < .001 (see Table 6). Previous academic achievement had the highest beta level (β = .537), followed by special education status (β = -.142), minority status (β = -.135), age (β = -.094), and sex (β = -.067). All predictors were significant at the .001 level except sex (p <.01). On average, males scored higher on the ACT than females, students in general education scored higher than those in special education, students with high previous academic performance scored higher than students with low previous academic achievement and students from non-minority backgrounds had higher test performance on the ACT than did students from minority backgrounds.

A second, post hoc analysis was conducted to determine the statistical significance of interaction terms with total test performance as measure by both ACT Composite scores and MME Social Studies performance. The same interaction terms were used from the analysis in Question 2, Part 1. The initial model included all of the 2x2 interaction terms. Similar to the first analyses, two general liner models were created that included the independent variables plus the interaction terms with two separate dependent variables (ACT Composite and MME Social Studies). Results indicated that there were no statistically significant interaction terms in either regression analysis. Given the non-significance of the interaction terms in analyses from both Part 1 and Part 2, there were no significant interaction terms to be included in a further post hoc analysis for Part 3 which included test anxiety as an independent variable.

Table 5

Stepwise Regression of Demographic Predictors of Test Anxiety

Model #[a]	Variable	Betas[b]	Std. Error	Sig.
1[c]	(constant)		.163	.000
	Age	.015	.046	.567
	Sex	.311	.052	.000
	Minority Status	-.061	.063	.026
	SES	.037	.059	.190
	Sped Status	.033	.081	.222
	GPA	-.124	.021	.000
2[d]	(constant)		.116	.000
	Sex	.310	.052	.000
	Minority Status	-.061	.063	.027
	SES	.037	.059	.184
	Sped Status	-.034	.081	.203
	GPA	-.125	.021	.000
3[e]	(constant)		.086	.000
	Sex	.310	.052	.000
	Minority Status	-.060	.063	.029
	SES	.036	.059	.204
	GPA	-.130	.020	.000
4[f]	(constant)		.064	.000
	Sex	.308	.051	.000
	Minority Status	-.069	.061	.010
	GPA	-.121	.020	.000

[a]Criteria for dropping predictor, $p > .100$.
[b]Betas are the standardized coefficients
[c]Full model
[d]Age dropped
[e]Special education status dropped
[f]Socio-economic status dropped

Table 6

Stepwise Regression of Demographic Predictors of Test Performance on ACT

Model #[a]	Variable	Betas[b]	Std. Error	Sig.
1[c]	(constant)		.742	.000
	Age	-.096	.211	.000
	Sex	-.065	.233	.007
	Minority Status	-.128	.310	.000
	SES	.036	.271	.166
	Sped Status	-.142	.401	.000
	GPA	.529	.092	.000
2[d]	(constant)		.704	.000
	Age	-.094	.211	.000
	Sex	-.067	.233	.006
	Minority Status	-.135	.303	.000
	Sped Status	-.142	.401	.000
	GPA	.537	.089	.000

[a]Criteria for dropping predictor, $p > .100$.
[b]Betas are the standardized coefficients
[c]Hypothesized model
[d]Final Model

In the second hypothesized model of test performance (MME Social Studies as outcome measure) the same demographic variables were included. Non-significant predictors were dropped (using $p > .10$) and no violations of multicollinearity or tolerance were found. In the initial model, socio-economic status was dropped as a predictor and in the second model age was dropped as a predictor resulting in a third, significant model (see Table 7). The final model predicted 29% ($R^2 = .285$) of the variance of test performance on the MME Social Studies subtest, $F(4, 1051) = 104.91$, $p < .001$. Despite similar predictors (other than age), the MME Social Studies model explained less variance in overall test performance than the ACT Composite model. Similar to the model predicting ACT Composite performance, previous academic achievement had the highest beta level ($\beta = .450$), but differed as sex ($\beta = -.209$), minority status ($\beta = -.126$), and special education status ($\beta = -.111$) were the next most important predictors. Regarding performance on the MME Social Studies subtest, on average, students with high previous academic achievement scored higher than students with lower previous academic achievement, males scored higher than females, students from non-minority backgrounds scored higher than students from minority backgrounds, and students in general education scored higher than students receiving special education services.

Table 7

Stepwise Regression of Demographic Predictors of Test Performance on MME Social Studies

Model #[a]	Variable	Betas[b]	Std. Error	Sig.
1[c]	(constant)		4.37	.000
	Age	-.041	1.25	.122
	Sex	-.210	1.38	.000
	Minority Status	-.118	1.83	.026
	SES	.041	1.60	.144
	Sped Status	-.110	2.37	.000
	GPA	.437	.543	.000
2[d]	(constant)		4.15	.000
	Age	-.039	1.25	.139
	Sex	-.212	1.38	.000
	Minority Status	-.126	1.79	.000
	Sped Status	-.110	2.37	.000
	GPA	.447	.527	.000
3[e]	(constant)		2.86	.000
	Sex	-.209	1.37	.000
	Minority Status	-.126	1.79	.000
	Sped Status	-.111	2.37	.000
	GPA	.450	.526	.000

[a]Criteria for dropping predictor, $p > .100$.

[b]Betas are the standardized coefficients

[c]Hypothesized model

[d]SES dropped

[e]Age dropped

Part 3—What is the relationship between test anxiety and test performance when controlling for

significant demographic variables?

A final hierarchical regression was conducted to examine the unique contribution of test

anxiety (as measured by FTAS total) in the explanation of test performance (as measured by

ACT Composite score and MME Social Studies score) when significant demographic predictors

(as identified in Part 1 and 2) were controlled. The author hypothesized that test anxiety would

be negatively related to test performance even after controlling for various demographic

variables. The variables were entered into the regression equation in two steps. Academic

achievement, special education status, minority status, and gender were entered into the first

regression block. FTAS total was entered into the second block. Results indicated no violations

of multicollinearity (VIFs; VIFs <10) and tolerance (Tols. > .10).

Results from step 1 indicated that variance in ACT performance accounted for by

previous academic achievement, minority status, special education status, and gender equaled .37

(R^2 = .37) which was significantly different from zero, F (4, 1042)= 155.11, p < .001. Academic

achievement was a significant predictor (β = .547, p < .001) as was special education status (β =

-.136, p < .001), minority status (β = -.123, p < .001) and gender (β = -.063, p < .05). After total

test anxiety was entered in step two, the change in variance accounted equaled .02 (ΔR^2) and

was significantly different than zero, F (5, 1041)= 130.87, p < .001. The second model

(including total test anxiety) accounted for 39% of the variance in ACT performance. Academic

achievement (β = .531, p < .001), special education status (β = -.128, p < .001), minority status (β

= -.128, p < .001) and test anxiety (β = -.120, p < .001) were significant predictors whereas

gender was not significant. Table 8 provides a summary of the model additions. On average,

students with high previous academic achievement scored higher on the ACT than students with

low previous academic achievement, students from non-minority backgrounds scored higher than students from minority backgrounds, students in general education scored higher than students receiving special education services, and students with lower levels of anxiety scored higher than students with higher levels of anxiety.

Table 8

Hierarchical Regression of Predictors of Test Performance on ACT Composite

Model #	Variable	Betas[a]	Std. Error	Sig.	Model R^2	ΔR^2
1[b]	(Constant)		.450	.000	.37	
	GPA	.547	.090	.000		
	Sped Status	-.136	.402	.000		
	Minority Status	-.123	.306	.000		
	Sex	-.063	.235	.011		
2[c]	(Constant)		.595	.000	.39	.02
	GPA	.531	.090	.000		
	Sped Status	-.128	.399	.000		
	Minority Status	-.128	.303	.000		
	Sex	-.023	.247	.370		
	FTAS Total	-.120	.127	.000		

[a]Betas are the standardized coefficients

[b]Initial model

[c]FTAS total added

A second, hierarchical regression analysis examined the relationship between test anxiety and test performance on the MME Social Studies subtest. Similar to the ACT analysis, academic achievement, special education status, minority status and gender were controlled for and entered into the first regression block. Test anxiety as measured by the FTAS total score was entered into the second regression block. No variations of multicollinearity were demonstrated.

Results from the first model (step1) indicated that the variance in MME performance accounted for by previous academic achievement, minority status, special education status, and gender equaled .29 (R^2 = .29) and was significantly different from zero, F (4, 1052)= 108.87, $p <$ .001. Significant predictors included academic achievement (β = .454, $p <$ .001), special education status (β = -.130, $p <$ .001), minority status (β = -.106, $p <$ .001) and gender (β = -.206, $p <$ .001; which was dissimilar from the ACT performance model). Similar to the ACT performance model, total test anxiety was entered into step two. The change in variance between step 1 and step 2 equaled .02 (ΔR^2) and was significantly different than zero, F (5, 1051)= 92.17, $p <$ .001. The second model accounted for 31% of the variance in MME performance. Academic achievement (β = .439, $p <$ .001), special education status (β = -.124, $p <$ .001), minority status (β = -.112, $p <$ .001) gender (β = -.168, $p <$ .001), and test anxiety (β = -.117, $p <$ .001) were significant. Table 9 provides a summary of model additions. On average, male students scored higher on the MME social studies than did females, students with higher previous academic achievement scored higher than students with lower previous academic achievement, students from non-minority backgrounds scored higher than students from minority backgrounds, students in general education classes scored higher than students receiving special education services, and students with lower levels of anxiety scored higher than did students with higher levels of anxiety as measured on the FTAS.

Table 9

Hierarchical Regression of Predictors of Test Performance on MME Social Studies

Model #	Variable	Betas[a]	Std. Error	Sig.	Model R^2	ΔR^2
1[b]	(Constant)		2.62	.000	.29	
	GPA	.454	.525	.000		
	Sped Status	-.130	2.25	.000		
	Minority Status	-.106	1.78	.000		
	Sex	-.206	1.37	.000		
2[c]	(Constant)		3.47	.000	.31	.02
	GPA	.439	.524	.000		
	Sped Status	-.124	2.24	.000		
	Minority Status	-.112	1.77	.000		
	Sex	-.168	1.44	.000		
	FTAS Total	-.117	.744	.000		

[a]Betas are the standardized coefficients

[b]Initial model

[c]FTAS total added

Question 3

Part 1—What is the relationship between student career goals and test anxiety?

A multivariate regression analysis was performed to examine the relationship between student career goals and test anxiety when significant demographic predictors (e.g., sex, minority status, and academic achievement) were included in the model. Career goals were aggregated into one of two groupings (1=4 year college, 0=2 year community college, vocational school, work force and military). Groupings were due to the hypothesis that students planning to attend a four year university had significantly different academic goals than students planning on entering the work force, military or community college. The initial model consisted of student career grouping as the independent variable, MME anxiety, ACT anxiety, and FTAS total as the dependent variables and sex, minority status and academic achievement entered as covariates.

There was a greater number of students planning to attend a four-year college. Descriptive statistics for anxiety levels between groupings is presented in Table 10. The Box Test of Equality of Covariance Matrices was significant indicating there was some variability across groups among the covariance matrices, $F (6, 1688173.129) = 4.876$, $p<.001$. However, large (and unequal) samples produce greater variances and covariances resulting in conservative probability values therefore significant findings can be trusted (Tabachnick & Fidell, 2007). The variability may be due to large, unequal sample sizes. Levene's test of equality of error variances was non-significant indicating that error variance of the dependent variable was equal across all three anxiety measures. The Wilks's Lambda was examined to determine if Ho = $\mu_1 = \mu_2$ which was nonsignificant with a p-value of 0.81. Total ACT anxiety and Total MME anxiety were significantly correlated ($r=.85$, $p<.01$). There was a significant difference among career groupings for Total ACT anxiety when controlling for sex, academic achievement and minority

66

status, with students going to four year universities reporting significantly higher anxiety on the

ACT, $F (1, 1246) = 7.07$, $p < .01$. There were no significant differences for Total MME anxiety,

$F (1, 1246) = 1.74$, $p = .19$ or FTAS $F (1, 1246) = .04$, $p = .84$.

Table 10

Descriptive Statistics for Anxiety Levels among Career Goal Groups

Anxiety Measure	Career Grouping	*M*	*SD*
FTAS Total[a]	4 Yr. College	69.42	22.53
	Non 4 Yr. College	69.91	22.20
	Total	69.53	22.45
Total MME Anxiety[b]	4 Yr. College	3.31	1.13
	Non 4 Yr. College	3.10	1.07
	Total	3.26	1.12
Total ACT Anxiety[c]	4 Yr. College	3.53	1.10
	Non 4 Yr. College	3.16	1.05
	Total	3.44	1.10

[a]FTAS Total calculated from the total sum survey responses.
[b]Total MME Anxiety calculated from survey items 45 and 55
[c]Total ACT Anxiety calculated from survey items 46 and 56

Question 4

A school administrator and test coordinator survey (see Appendix J) was administered at each of the six schools; five of the questionnaires were returned. All five respondents reported they were anxious for the upcoming test and thought that the majority of their staff and students were also anxious for the upcoming test. Also, all of the schools engaged in test preparation activities; these activities included one school with an ACT prep course (for "college bound" students), another reported having a "test prep assembly", and the three remaining schools engaged in general test reminders (e.g., announcing the test on the morning report, reminding students to inform their parents about the upcoming test). Test administration procedures among schools were similar with four schools reporting administering the test in the gymnasium or lunch room and one school reporting the test administered in individual classrooms.   One school administrator indicated that students in his school (that made AYP) were proud of their test results, whereas another administrator (in a school that did not meet AYP) stated that school staff and students did not expect to perform well on the upcoming test. None of the test coordinators or school administrators reported that their staff engaged in formal test anxiety interventions. While there were some differences (e.g, ACT prep course), the majority of school wide test preparation activities were similar among the respondents.

Part 1—What is the relationship between school AYP status and teacher anxiety?

An independent samples t-test was conducted to determine if there were significant differences in total teacher reported anxiety by school AYP status. The Levene's test indicated that the homogeneity of variance assumption is tenable ($F >.05$), so equality of variances can be assumed. The mean anxiety scores of teachers in schools that have made AYP ($M=7.42$, $SD=2.29$) were lower than teachers in schools that had not made AYP ($M=8.15$, $SD=2.42$).

However, results indicate that there was no significant difference in teacher anxiety among schools that have made AYP and schools that have not made AYP, t (116)= -1.67, p=.098.

Summary and descriptive statistics for teacher related anxiety and attitudes regarding high-stakes tests are presented in Figures 2, 3, and 4. Teacher reported anxiety (N=118) was examined for descriptive purposes. Over 40% of teachers indicated somewhat agree or strongly agree that they are anxious for their students to take the upcoming high-stakes test. When asked about their colleagues, 44% of teachers indicated their belief that the majority of teachers within their respective schools were anxious for the upcoming test. Moreover, data indicated 83% of teachers believed that administrators at their schools reported high levels of anxiety regarding the upcoming the MME/ACT.

Part 2—What is the relationship between school AYP status and student anxiety?

An exploratory analysis was conducted to examine the relationship among school AYP status and student anxiety. Descriptive statistics for student anxiety by school are presented in Table 11. A MANOVA was used to analyze ACT anxiety and total FTAS anxiety in students in schools that made AYP and students in schools that did not make AYP.

Results from the Box's Test of Equality of Covariance Matrices were non-significant indicating covariance matrices were equal across groups, F (3, 3.462E8) = 1.812, p= .142. Next, the Wilks's Lambda was examined and found to be significant ($p < .01$) indicating group mean difference among total FTAS anxiety and ACT anxiety. Data indicate a significant difference among students in schools that made AYP and those that have not for total anxiety on the FTAS, F (1, 1298) = 6.14, $p < .05$, and total ACT anxiety F (1, 1298) = 10.92, $p < .01$. Examining mean differences (as indicated in Table 11.) suggests that students in schools that have consistently

made AYP have higher test anxiety as measured by the FTAS total and ACT anxiety than students in schools that have consistently not made AYP.

Part 3—What is the relationship between teacher and student anxiety?

A bivariate correlation was used to analyze the relationship between teacher anxiety (from a principal component analysis) and student anxiety. If this correlation was significant, then the variables would be analyzed using a multiple regression with the significant demographic predictors entered along with teacher anxiety to predict student anxiety. The predicted association between student anxiety and teacher anxiety was found to be nonsignificant, $r(1243)= .04, p > .05$.
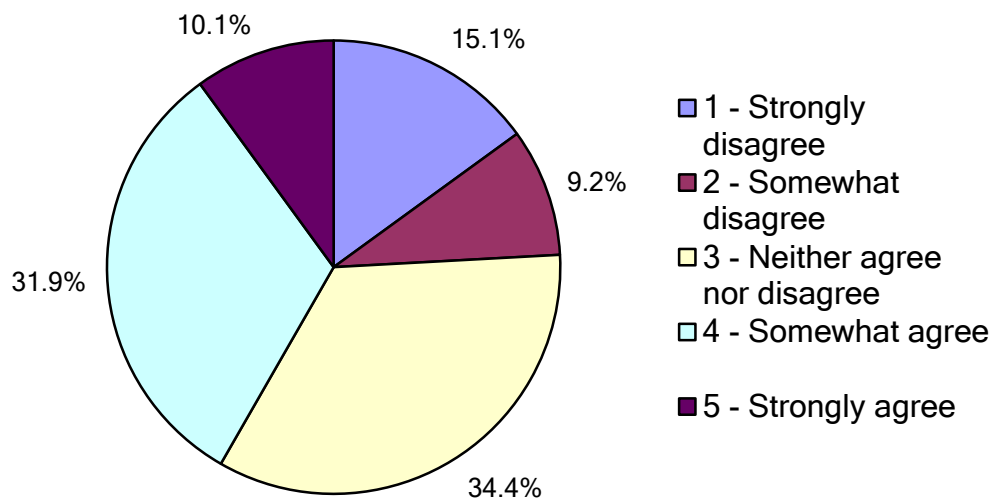
Figure 2. Teacher Reported Anxiety for Students to take the MME/ACT

For interpretation of the references to color in this and all other figures, the reader is referred to

the electronic version of this dissertation.

5.9%   1.7%

18.6%

38.1%

35.6%

1 - Strongly disagree

2 - Somewhat disagree

3 - Neither agree nor disagree

4 - Somewhat agree

5 - Strongly agree

The majority of teachers in my school are anxious about the MME/ACT.
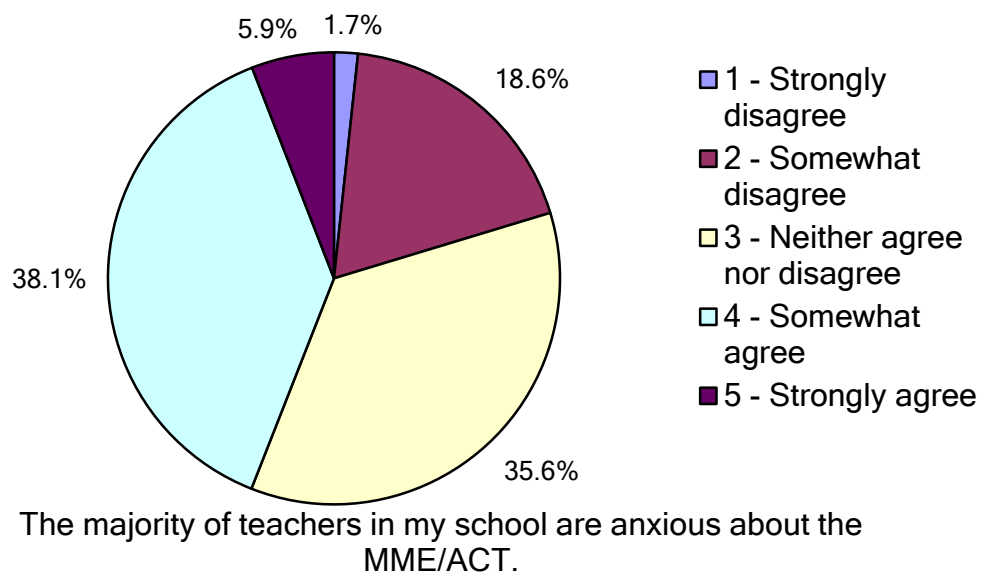
Figure 3. Teacher Perception of Peer Anxiety

The majority of administrators in my school are anxious about the MME/ACT.

Legend:
- 1 - Strongly disagree
- 2 - Somewhat disagree
- 3 - Neither agree nor disagree
- 4 - Somewhat agree
- 5 - Strongly agree

0.8% 4.2% 11.0% 40.7% 43.2%
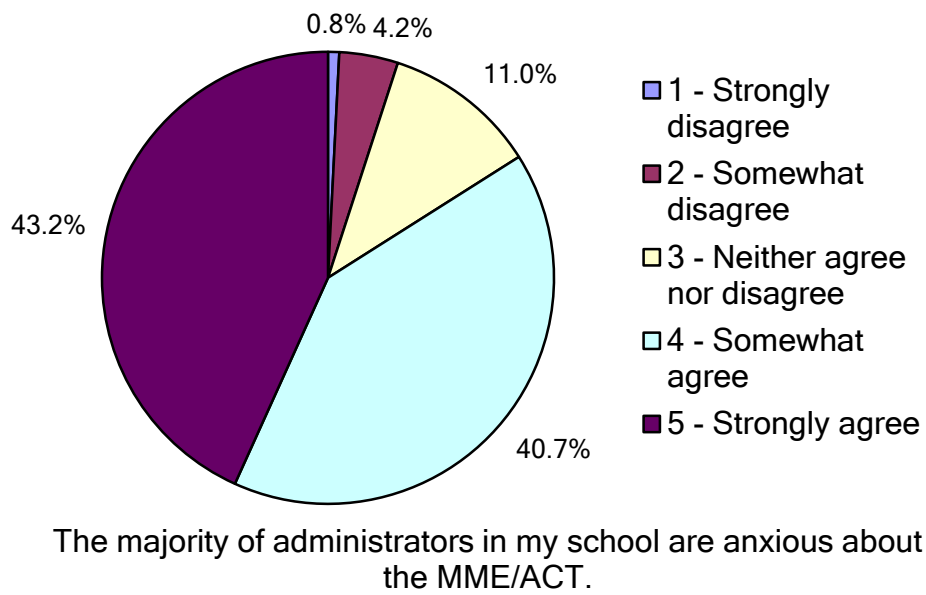
Figure 4. Teacher Beliefs about Administrator Anxiety

Table 11

Student Anxiety levels by AYP status

School	AYP Status	FTAS Total		ACT Anxiety		
		M	*SD*	*M*	*SD*	N
A	Met	69.23	22.54	3.47	1.10	235
B	Met	73.60	24.15	3.49	0.99	126
C	Met	71.30	23.69	3.61	1.21	268
Average	Met	71.07	23.46	3.53	1.13	629
D	Identified for Restructuring	67.85	20.93	3.39	1.08	371
E	Identified for Restructuring	68.08	23.92	3.18	1.07	182
F	Identified for Restructuring	67.84	21.62	3.35	1.03	117
Average	Not Met	67.84	21.85	3.33	1.07	670

CHAPTER 5

DISCUSSION

Education in the United States has fundamentally changed since the enactment of accountability policy (e.g., NCLB) through the use of high-stakes tests. Test-based accountability policy has been associated with a wide range of consequences, yet little is known about student and teacher anxiety levels within these high-stakes situations. There have been very few research studies which have examine the relationship of test anxiety and high-stakes tests since the implementation of NCLB (Segool, 2009). Due to the enormous weight placed upon high-stakes test scores by government, schools, teachers, administrators and parents, it is crucial to understand the relationship between test anxiety and test performance. The purpose of this study was to examine the student and teacher experience within this environment as it relates to an ecological model (See Figure 1). A critical analysis took place of temporary, state anxiety versus longstanding, trait anxiety prior to the Michigan Merit Exam. The differences (or lack thereof) between trait and state anxiety provide important insights into the nature of test anxiety before high-stakes examinations. Demographic predictors of both test anxiety and test performance were also analyzed. While there has been some previous evidence to suggest relevant predictors of test anxiety, there has been little published research (Segool, 2009) in the US on student test anxiety since the enactment of NCLB, specifically with high-stakes tests for high school students.

Broad generalizations are often made about the consequences and implications of high-stakes tests. However, students may consider the stakes differently depending on their individual career goals. This study examined the individual stakes of an exam by comparing career goals and the manifestation of test anxiety. This information could provide a more nuanced

understanding of the perceived nature of test outcomes or "stakes" of a test for high school students.

The influence of external accountability (i.e., government designated performance goals such as AYP) with respect to student and teacher anxiety was explored. Potential differences in teacher and student anxiety were examined in schools that have consistently made AYP and schools that have not made AYP for a minimum of three consecutive years. This study expands upon the current test anxiety literature by 1) directly examining temporary anxiety levels prior to a high-stakes exam, 2) identifying significant predictors of test anxiety and test performance within a high-stakes situation, 3) examining the difference between career groups with respect to ACT anxiety and 4) exploring the contribution of external accountability with the manifestation of anxiety.

State versus Trait Anxiety

Differences among state and trait anxiety were examined within this study. Previous research suggested that state and trait anxiety would be similar under "normal" conditions as measured on the STAI (Joesting, 1975). It was hypothesized that state anxiety would be significantly greater than trait anxiety one week prior to the Michigan Merit Exam. This hypothesis was not supported, as trait anxiety was actually significantly higher than state anxiety. This finding was different from previous research and may be due to several reasons. Students may not have been anxious one week prior to the MME (or at least not so much as to be different than typical anxiety levels). Research has suggested that student anxiety levels increased several days before an examination (Rafferty, Smith & Ptacek, 1997). Given the heterogeneous nature in timing of survey administration (i.e., one school administered the survey one week prior, another school administered 3 days prior), results may have been compromised as it relates specifically

to state type anxiety. Alternatively, schools often make students aware of the upcoming high-stakes test early in the school year and then provide reminders on a regular basis. It is possible that students have a consistent level of anxiety associated with the high-stakes test for a much longer period of time than one week before the test or students have become "desensitized" to the anxiety specifically associated with the upcoming test. Qualitative information from the test coordinators indicated a school wide test preparation activity (e.g., holding a rally or an assembly) at one school whereas the other coordinators reported "general reminders" delivered by teachers and/or a letter sent home to parents. Future studies should administer the STAI on the same day (or at least the same number of days prior to a high-stakes test) in addition to a systematic analysis of test preparation activities for more accurate anxiety level comparisons. In addition, previous research (Rafferty, Smith & Ptacek, 1997) used a qualitative method (e.g., anxiety diaries and interviews) and a small sample whereas the present study included a much larger sample and quantitative comparisons that utilized an established anxiety scale (e.g., STAI); therefore, results from this study may be more reflective of student anxiety levels before an upcoming high-stakes test.

One methodological concern was that the STAI portion of the survey was the last 40 questions of a 96 question survey. Some participants may have suffered from response fatigue and did not accurately complete the STAI. Out of a total 1465 participants, 1161 completed the STAI while over 300 failed to complete the full survey. The directions on both parts of the STAI were subtly different (see Appendix C). This also may have contributed to a less than accurate response. Future research should a) administer a brief version of the STAI (Marteau & Bekker, 1993) as part of the total survey thus reducing the number of survey items, b) administer the STAI at a separate point in time, c) rely solely on other measures of anxiety (e.g., FTAS) and d)

randomize the order of the anxiety instruments thus counterbalancing for the effect of student fatigue on one part of the survey.

Even though the proposed directional difference among state and trait was not supported, it is interesting to note the percentile ranks for the *average* (raw) state and trait anxiety scores for males were at the 54th and 56th percentiles and females at the 69th and 68th percentile. As indicated in the result section, mean scores for trait and state anxiety were compared to the percentile rank as indicated in the STAI manual. These data suggest a number of students were indicating generally high levels of anxiety (both state and trait) as measured on the STAI. While there may not have been significantly higher state versus trait anxiety, anxiety levels as a whole were somewhat high. The State median was the same as the State mean whereas the Trait median (*Mdn*=44.00) was slightly higher than the trait mean (*M*=43.02) thus the possibility exists that a small number of students with low anxiety may be depressing the overall Trait mean and the overall number of students with high anxiety may be greater than expected. These data, similar to the mean levels of state and trait anxiety suffer from the same limitations as noted above.

The STAI was last normed in 1983. It could be argued that the testing climate in education today is significantly different than in 1983, particularly since the passage of NCLB. Therefore, students may be more anxious than students were nearly 30 years ago. This proposition may explain why average raw scores for males and females were higher than what would be predicted. The STAI may not accurately capture student anxiety in today's high-stakes educational climate. Future research should consider using anxiety assessments with more recent norming data.

Demographic Predictors of Test Anxiety

This study examined the relationship among student demographic variables and test anxiety. A meta-analysis conducted by Hembree (1988) identified several significant demographic variables associated with the manifestation of test anxiety including gender, age, ethnicity, and socioeconomic status. Other studies have identified academic achievement (King & Ollendick, 1989) and special education status (Sena, et al., 2007) as significant predictors of test anxiety. A stepwise regression identified minority status, academic achievement, and sex as significant predictors of test anxiety (i.e., total test anxiety on the FTAS). This finding only accounted for 10% of the variance in test anxiety thus demographic variables alone may not be sufficient to predict test anxiety.

Special education status, socioeconomic status, and age were not significant predictors of test anxiety and dissimilar to previous studies (Sena, et al., 2007, Putwain, 2007, 2008; Hembree, 1988). Contrary to the conclusion of Sena and colleagues (2007), special education status did not predict the presence of test anxiety. Also, students with severe disabilities (e.g., multiple handicapped, cognitively impaired) were not included in the present study, including those who attended alternative education programs, who qualified for the alternate assessment, or could not read the survey materials. Therefore, a number of potential participants who received special education services were not included in the sample (as indicated by the aforementioned reasons). If these students (i.e., those with severe disabilities) were included in the present study, the relationship of special education status and test anxiety may have been different. Age was most likely a non-significant variable as there was little age variability among the eleventh grade participants. Comparing test anxiety levels among ages may be more relevant when also including students from different grades. Regarding socioeconomic status, this study grouped

students as those receiving free or reduced lunch and those who do not. Previous research (Putwain, 2007) had indicated multiple levels of SES based upon parental employment/income rather than the two levels in this study. Arguably, this study provides a clearer distinction among SES as students qualifying for free or reduced lunch must meet strict poverty guidelines as opposed to an employment analysis which may reveal great variability in potential income (and does not account for other extenuating variables such as family size).

The finding of Caucasian students reporting higher levels of test anxiety than students from minority backgrounds was contrary to previous research (Turner, et al., 1993; Beidel, et al., 1994; Putwain, 2007). Previous studies which had used relatively small samples (N=62, Beidel, et al., 1994) and examined anxiety levels specific to African-American children (Turner, et al., 1993) whereas this study included in the "minority" classification students from Asian (N=49), African-American (N=69), Hispanic (N=94), Native American (N=21) and Multi-racial (N=124) backgrounds. A more nuanced analysis among racial categories may reveal differences among minority backgrounds; however, this group (i.e., not Caucasian) taken as a whole was less anxious than their Caucasian peers. Another possible explanation for this difference was the inclusion of multiple control variables (e.g., gender, SES, academic achievement) that may not have been present in previous studies.

Gender was found to be a significant predictor of test anxiety, with females reporting higher levels of test anxiety than males. This finding is consistent with many previous studies (Hembree, 1988, Putwain, 2007, 2008; Lowe & Lee, 2008) indicating higher levels of test anxiety within females. Additionally, academic achievement was identified as a significant predictor of test anxiety. Previous research has suggested that there is an inverse relationship among GPA and test anxiety levels (Speilberger, 1966, Hembree, 1988); results from this study

81

supported this notion as data indicated higher levels of test anxiety in students with lower reported GPA. However, student self-reported GPA was a methodological limitation of this study. Self report may have led to an inflation of GPA thus complicating the analysis of academic achievement and test anxiety. In addition, GPA is calculated from student grades in a variety of classes and may not be indicative of academic ability in specific domains (i.e., GPA includes classes such as physical education and choir which may inflate/deflate GPA as a predictor of academic achievement in math or science). Schools often have different curricula thus complicating the comparison of student GPAs in different schools. A better indicator of academic achievement may include previous test performance (e.g., performance on the Michigan Educational Assessment Program or, in earlier grades, performance on the Iowa Test of Basic Skills). The analysis of student career goals may provide more insight into this relationship with external motivators or goals (e.g., desire to attend a four year college) than GPA.

<div align="center">Demographic Predictors of Test Performance</div>

Several significant demographic predictors of test performance were identified within this study. After comparing the results of stepwise regression analysis of predictors of both ACT performance and performance on the MME Social Studies subtest, gender, minority status, special education status, and previous academic achievement were identified as significant. These results are consistent with previous research studies examining high-stakes test performance (Putwain, 2008a, 2008b). Similar to the prediction models of test anxiety, previous academic achievement accounted for the largest amount of variance.

A moderate amount of the variance in overall test performance on the ACT was accounted for by demographic variables ($r=.39$), whereas demographic variables accounted for a

smaller amount of variance in MME Social Studies performance ($r=.29$). The difference in variance accounted for may be due to several causes. First, students may be more inclined to take the ACT seriously as data indicate students reporting higher anxiety levels on the ACT as opposed to the MME. Therefore, the model may be more precise in predicting ACT performance with a smaller percentage of variance due to unexplained or random factors than performance on the MME Social Studies subtest. Second, fatigue may have been a factor as students took several tests over the course of three days, with the ACT administered earlier than other subsections of the MME (including the Social Studies subtest). Finally, while previous research (Putwain, 2008) has identified significant demographic predictors of *overall* academic performance, there may be a unique factor specific to Social Studies type content which was not accounted for within the prediction model.

<center>Relationship between Test Anxiety and Test Performance</center>

This study examined the influence of test anxiety, while controlling for significant demographic variables, on student test performance on the ACT Composite and the MME Social Studies subtest. A hierarchical regression analysis was used to examine the contributing variance of significant demographic predictors (as identified in the previous two analyses) and test anxiety as measured by the FTAS total score. Test anxiety accounted for 2% of the total variance (39%) in test performance on the ACT when controlling for significant demographic variables such as special education status (2%) and minority status (1%). On the MME Social Studies test, test anxiety accounted for 2% of the total variance (31%) in test performance. Previous academic achievement accounted for the largest amount of variance in the predictor model on the ACT (33%) and MME Social Studies test (22%). Gender was a non-significant variable in the ACT prediction model and significant in the MME Social Studies model.

<center>83</center>

Results from this study support the hypothesis that test anxiety is a significant (negative) predictor of test performance. This result was consistent with previous studies (Putwain, 2008) that found a similar percentage (2-7%) of variance in test performance accounted for by test anxiety when controlling for demographic variables. While accounting for a relatively small percentage of variance in overall test performance, test anxiety may be considered an area for intervention in highly anxious groups of students as schools are facing increasing difficulties in meeting AYP. It is important to note that some schools focus resources on groups of students who are close to passing the test or moving into the next level of achievement category at the expense of improving test performance of all students (Diamond & Spillane, 2004). Therefore, 2% of the variance in test performance accounted for by test anxiety may be significant for those schools attempting to increase the performance of students who may be close to passing the test, and reducing test anxiety may be enough to achieve this goal. This study did not specifically identify how many students may have been able to move from one achievement category to the next based upon an assumed reduction in test anxiety and subsequent improvement in test performance.

Decisions about implementing interventions for test anxiety should be considered relative to the potential variance accounted for by other variables (i.e., previous academic performance accounted for a much larger percentage of the variance in test performance) and potential cost (e.g., in time, dollars spent and at the expense of implementing other interventions such as a study skills curricula). Additional research is needed to identify clinical levels of test anxiety on the FTAS, and if these groups (e.g., "highly test anxious" versus "low test anxious") differ in test performance.

Relationship between Test Anxiety and Career Goals

Previous research regarding the relationship of high-stakes tests and test anxiety has typically assumed external consequences are reflective of the actual "stakes" of the examination (Putwain, 2007, 2008). This study is unique in that it addresses the individual nature of "stakes" and its relationship with test anxiety by identifying student career goals. Students were grouped into those indicating a desire to attend a four year university (N=964) and those wishing to attend community college, military service or enter the workforce (N=267). When controlling for academic achievement, sex, and minority status, students in the four year university group reported significantly higher anxiety on the ACT than those in the non four-year university group. Although the pattern was similar for anxiety on the MME, significant differences were not identified between these career groups.

Ryan and colleagues (2007), in a small scale (N=33) qualitative study indicated that career goals influenced the manifestation of test anxiety. In other words, students who had high career aspirations (e.g., four year university) were more likely to report higher levels of anxiety for an upcoming test. The results of this study are consistent with the aforementioned research in that students wishing to attend four year universities indicated higher ACT type anxiety. However, there were no differences on MME anxiety. This may be due to the relative importance placed upon ACT test scores by college bound students.

Finally, the consequences assumed from the implementation of high-stakes tests may not accurately capture experience of individual students. Because there are different levels of test anxiety between students with different career aspirations, the stakes of the test cannot be fully explained by external consequences (or those assumed by NCLB). This study is the first to ask students about their own, *individual*, perceptions of test consequences with regards to career

aspirations and test anxiety. Previous research has assumed "stakes" given the consequences attached to test performance on a state wide level (e.g., graduation tests in several states) and a test wide level (i.e., assumption that each day and/or part of a test has high stakes for students). However, data from this study indicate heterogeneity in "stakes" of a test given the differing anxiety levels as it relates to separate test components. Future research should be careful to assume universality of test stakes due to the varying nature of student career goals and the relevance of test performance as a means to those ends. These data add to our understanding of relevant predictors of the manifestation of test anxiety and which groups of students may be more susceptible to test anxiety than others.

It should be noted that there is limited technical support for the use of the MME anxiety and ACT anxiety measures. Given that these variables were created for the present study, there is a general lack of available technical adequacy information. The MME and ACT anxiety variables only included two survey items a piece (e.g., items 45, 46, 55, 56 on Appendix C). The item correlations were considered weak (Cohen, 1977), yet significant. Due to these limitations, caution must be exercised when interpreting the use of these two indicators as differences in anxiety levels among career groups.

<center>School AYP Status and Student and Teacher Anxiety</center>

Brofennbrenner's (1979) ecological systems theory was the organizing conceptual framework used in this study; it was hypothesized that government accountability targets (e.g., AYP) would influence teacher anxiety thereby influencing student anxiety specifically in schools that have consistently failed to meet AYP (See Figure 1). Practical limitations (i.e., lack of participating schools) prevented a systematic exploration of this theory. To do so, many more schools would need to be included (and randomly selected) to examine the influence of external

<center>86</center>

accountability on levels of teacher and student anxiety. Differences in teacher and/or student anxiety among schools may especially be important if inordinate numbers of teachers and students were anxious in schools that failed to make AYP; given the significant negative relationship of test anxiety with test performance identified in this study, these schools may face a unique hurdle to make AYP. Therefore, an exploratory analysis was conducted between groups of student and teachers in schools that have consistently made AYP and schools that have not made AYP for at least three consecutive years.

Pedulla and colleagues (2003) identified high levels of teacher anxiety correlated strongly with the stakes of the test, with the higher the stakes the higher the levels of teacher anxiety. Their study also indicated higher levels of student anxiety in states that used high-stakes tests. However, researchers should be cautioned against generalizing the findings of the influence of high-stakes tests and accountability policy due to the great variation within and between states in their design and implementation of test-based accountability (Ysseldyke, et al., 2004). Arguably, an examination of AYP among schools (or teachers within those schools) would provide a more nuanced analysis of the potential influences of AYP because an upcoming test may be more important and carry more severe consequences for schools that have not made AYP for several years.

Dissimilar to the Pedulla study, data from this study revealed no significant differences among teachers in schools categorized as "high-stakes" (i.e., those schools that had not made AYP) versus "low-stakes" schools (i.e., schools that have consistently made AYP). This study is different from Pedulla in that teacher anxiety was quantitatively measured within schools of specific AYP status rather than a multiple state survey (i.e. schools of different AYP status were aggregated and then compared as groups between states in the Pedulla study). This may explain

the different results from this study versus the Pedulla study regarding teacher anxiety in high-stakes environments. A failure of one high-stakes school to administer the teacher survey may have compromised results although mean anxiety scores were higher for teachers in high-stakes schools. Despite no significant differences between schools, teacher anxiety on average was relatively high. Even in schools that had consistently made AYP, teachers reported feeling anxious for the upcoming test. Descriptive statistics indicated over 40% of teachers reported that they were anxious or somewhat anxious for the upcoming test and 83% believed administrators were anxious for the upcoming test. As proposed within the conceptual framework (Figure 1), there may be external variables (e.g., administrator pressures, the use of student test performance data for teacher evaluations) other than AYP which may be influencing the manifestation of teacher anxiety. These data suggest more analysis is needed to evaluate the influence of external accountability policy on the levels of teacher anxiety.

When levels of student anxiety were examined, results indicated higher student anxiety (as measured on the FTAS total and ACT type anxiety) in schools that have consistently *made* AYP. Just because a school has made AYP does not automatically indicate that students in that school will have less or even lowered test anxiety than if that school had not made AYP. This result was opposite from the proposed hypothesis. Students in schools that have consistently made AYP may be more cognizant of the importance of the upcoming test thus resulting in higher levels of anxiety and the consistent pressure to continually score higher and higher (based upon increasing performance targets needed to meet AYP and eventually reach 100% proficiency). This is partially supported by interviews from administrators and test coordinators. One administrator (from a school that had not made AYP) indicated that students did not care about the upcoming test whereas another administrator (from a school that had made AYP)

reported that students were highly invested in the test because most students viewed the ACT portion of the MME as necessary for college admission. It should be noted that test anxiety (i.e., arousal) in small amounts may be useful for test performance as originally suggested by Yerkes and Dodson (1908). Indeed, in fields such as kinesiology and social psychology, research has suggested that arousal (i.e., anxiety) does not decrease performance (Thibodeau, Gomez-Perez, & Asmundson, 2012; Barnard, Broman-Fulks, Michael, Webb, & Zawilinski, 2010) and in some instances, such as practice Graduate Record Examinations, has predicted improved performance (Jamieson, Mendes, Blackstock, & Schmader, 2009) in comparison to low levels of arousal. However, test anxiety research has assumed a negative linear relationship with test performance since the 1950s (Mandler & Sarason, 1952; S. B. Sarason & Mandler, 1952; Spielberger, 1966; Liebert & Morris, 1967; Zeidner, 1998; Sena, et al., 2007; Putwain, 2008b). Recent research studies such as Segool (2009) suggested that students with low levels of test anxiety perform higher on tests than do students reporting medium and high levels of test anxiety. More research and analyses are needed (e.g., curvilinear analyses) to examine the relationship between test anxiety and test performance to identify an optimal level of test anxiety or arousal.

A final analysis examined the relationship of student and teacher anxiety. As posited by Doyle and Forsyth (1973), teacher anxiety was hypothesized to have a significant positive correlation with student test anxiety (i.e., higher teacher anxiety would result in higher student anxiety). Furthermore, the conceptual framework (See Figure 1) suggested that there were multiple levels of influence on the manifestation of student test anxiety; this includes macro system variables to meso level systems. Results from this study did not support the proposed relationship of teacher and student anxiety. There are several reasons why this hypothesis may have not been supported. A methodological limitation of this study prevented matching of

89

individual teachers with student participants. Therefore, teacher anxiety was aggregated and then assigned to each student within that respective school. This method assumes 1) a teacher aggregate score was truly reflective of total anxiety within that school and 2) each student in the school interacted with at least one teacher from the aggregated data A second limitation involved the sampling of teachers. Roughly 30% of teachers at each school responded to the survey compared to an average student response rate over 80%. It was not known whether the included teacher sample was reflective of the overall anxiety among the school staff

Limitations

There were several limitations within the present study beyond what was previously noted. Six high schools agreed to participate in this research. While these schools represented a wide range of settings, geographic areas within Michigan, and student demographics, they may not be representative of Michigan or the general population at large. Also, the administration and consequences of high-stakes tests are vastly different among states within the US, with each state department of education having the flexibility to determine yearly target scores and some variation in sanctions. Therefore, it would be unwise to draw conclusions from this study regarding the relationship of test anxiety and test performance on high-stakes tests in other states.

Student participation rates ranged from 53% to 88% which is considered excellent for survey research (Hopkins & Gullickson, 1992). There were several students who did not fully complete the survey and whose data was therefore not interpretable. Therefore, the participating sample may not be representative of the school. For example, students in alternative schools and students who took the alternate assessment (e.g., the "2%" assessment as indicated by NCLB) were not included in the sample, thus potentially missing out on important student subgroups such as those with cognitive impairments or behavioral difficulties. In addition, the mode of

survey administration could have potentially varied among schools. The researcher was only able to personally administer surveys at one school, while the other schools opted to administer in 1) a large gymnasium by the school principal/test coordinator or 2) individual classrooms by teachers. While standardized directions were provided, there remains a possibility of variation in the survey administration language and how the survey was explained to students.

Teacher anxiety levels should be interpreted with caution. One school (out of six) chose not to participate in the teacher surveys. The remaining five schools had response rates averaging 25% of the total teacher population. It is unknown whether participants were representative of teacher anxiety levels of the whole school. Also, question four is intended to be exploratory in that a small sample of schools (N=6) was not sufficient to determine significant differences in teacher or student anxiety levels with respect to AYP.

Similar to the teacher survey, the same school chose not to participate in the matching of participant test scores and test anxiety scores. Thus, over 200 participants were dropped from the analysis of the relationship between test anxiety and test performance. The addition of these excluded participants could possibly have changed the relationship or the percentage of variance accounted for by test anxiety. It should be noted that the sample used in Question 3, part 2 and 3 was much larger (N=1046) than the amount required by a power analysis to determine a medium effect size.

<div align="center">Future Research</div>

Future research could address several of the limitations within the present study. More sophisticated analyses (e.g., Structural Equation Models or SEM) may better account for the multitude of continuous and categorical variables identified within this study than regression analyses, thus reducing the probability of Type I errors. These analyses may also include

<div align="center">91</div>

multiple measures of test anxiety (e.g., FTAS and STAI) and/or subtests of broader measures (e.g., Social Derogation on the FTAS), leading to an improved understanding of the unique influence of different aspects of test anxiety with test performance.

SEM analysis could also address a major limitation of using the FTAS, the lack of established clinical levels. As currently addressed, the FTAS scores and test performance are treated as continuous variables. Latent profile analysis (LPA) may be used by researchers in future studies to examine the presence of dichotomous groups of test anxious individuals (i.e., highly anxious, moderately anxious on the FTAS). Descriptive statistics suggest a small number of individuals with high scores on the FTAS. If categorically identified with LPA, data from this study may help to identify clinical levels of test anxiety as measured on the FTAS.

As the test coordinator and school administrator interviews suggest, some teachers and students may have given up or did not expect to perform well on the high-stakes test thus complicating the relationship between test anxiety and test performance (i.e., very high or low anxious individuals may perform poorly due to lack of efficacy, motivation, etc). Additional research needs to address measurement issues such as the differing nature of test anxiety as it relates to specific tests and a critical examination of the biopsychosocial model (Lowe, et al., 2008) as a viable interpretation of the test anxiety phenomenon.

Implications

This study explored how teachers and students perceive a "high-stakes test environment" by examining anxiety levels as related to demographic predictors, test performance, school AYP status and student career goals. These relationships were examined within an ecological framework based upon Bronfenbrenner's (1979) ecological systems theory which posits that there are multiple levels of systems which have direct influences on each other. As applied to

this study, this framework was unique to previous test anxiety research (Putwain, 2007, 2008; Gregor, 2005) that did not incorporate "social" levels of influence (e.g., pressures from peers, parents, teachers) The present study utilized a test anxiety measure (FTAS, Friedman & Bendas-Jacob, 1997) that included anxiety from social sources such as parents and teachers. Results from this study have several important implications for school psychologists and educators. First, this study provides a more nuanced understanding of the test anxiety phenomenon, specifically manifested in a high-stakes environment, by identifying demographic predictors along with student career aspirations. This was the first study to identify individual stakes by examining the relationship of test anxiety with career aspirations. Moreover, depending on their career aspirations, there was a significant difference among career groups for ACT anxiety. This knowledge is valuable to educators in preparing different groups of students before specific tests depending on their career goals and aspirations. For instance, a college-bound student may be more anxious for the ACT portion of the high-stakes test than a student interested in entering the workforce. With limited resources, schools are often forced to prioritize interventions (Diamond & Spillane, 2004). This study provides a clearer picture of which groups of students may be the most susceptible to test anxiety based on demographic characteristics and career aspirations.

While there was no significant difference in teacher anxiety by school AYP status, approximately 42% of teachers in the sample indicated that they strongly agree or somewhat agree that they are anxious for the upcoming test (compared to 24% who indicated strongly disagree or somewhat disagree and 34% who indicated neither agree nor disagree). When asked if their fellow teachers were anxious for the upcoming test, 44% of respondents indicated strongly agree or somewhat agree (compared to 20% who indicated strongly disagree or somewhat disagree and 36% who indicated neither agree nor disagree). Additionally, 83% of

93

participants indicated strong agreement or somewhat agreement that school administrators were anxious for the upcoming test. As interpreted though Bronfenbrenner's (1979) ecological model and the conceptual framework used in this study, teacher anxiety may influence the manifestation of student test anxiety. However, student anxiety was actually higher in schools that made AYP. There may be additional ecological variables, other than teacher anxiety, that were unaccounted for in this study and may be influencing the manifestation of test anxiety.

With the increased use of large-scale standardized test outcomes to make high-stakes decisions regarding graduation and promotion for students, teacher evaluations for tenure and effectiveness ratings assigned to schools, so too does the pressure surrounding the testing environment. Data from this study indicate between 2-3% of the variance of test performance accounted for by test anxiety for *all* students. However, these data do not identify "clinical" levels of test anxiety (i.e. groups of students with high levels of anxiety and thus low levels of test performance). When different levels of test anxiety have been identified, students with low test anxiety perform significantly greater than do students with medium and high levels of anxiety (Segool, 2009). Additionally, 25% of students were afflicted with high levels of test anxiety in a recent study (Bradley, et al., 2007). Schools are largely remiss in teaching children the skills necessary to understand and self-regulate the emotional stress and anxiety associated with testing (Greenberg, et al., 2003; Mayer, et al., 2008), and there have only been four test anxiety treatment studies conducted in U.S. public schools published over the past decade. Students suffering from high levels of test anxiety perform poorly on tests (Hembree, 1988; McDonald, 2001), which may result in underestimates of student achievement and school effectiveness. School psychologists can serve an integral role in assessing variables (including test anxiety) which may be suppressing student test performance. They can also serve as

94

resources for schools wanting to prevent and treat test anxiety at multiple levels of service (e.g., school wide initiatives versus individual treatment).

School psychologists can serve as leaders in the assessment and treatment of test anxiety. For schools which may have a large portion of students with high anxiety, Weems and colleagues (2010) have provided a detailed and thorough description of a test anxiety prevention and intervention program through the University of New Orleans. Freely available test anxiety assessments such as the FRIEDBEN Test Anxiety Scale (Friedman & Bendas-Jacob, 1997) and Children's Test Anxiety Scale (Wren & Benson, 2004) can be used to screen targeted groups of students with relative ease and minimal intrusiveness. With data in hand, school psychologists can assist schools in identifying targeted groups or highly anxious individual students for intervention support. Systematic reviews have identified effective interventions in reducing test anxiety for different groups of students (Ergene, 2003; von der Embse, Barterian & Segool, in press). At the group level, multi-method cognitive-behavioral interventions (Gregor, 2005) or more specific behavioral (Egbochuku & Obodo, 2005; Larson, Ramahi, Conn, Estes & Gibellini, 2010), cognitive (Lang & Lang, 2010), or academic interventions (Carter et. al, 2005; Faber, 2010) can be delivered to targeted classrooms or groups of students with high levels of test anxiety who have not responded to universal prevention and intervention efforts. At the most intensive individual level of service, relaxation training using biofeedback software may be used to teach physiological self-control and to evaluate intervention effectiveness for severely test anxious individuals (Bradley, et al., 2010; Yahav & Cohen, 2008).

APPENDICES

Appendix A.

FRIEDBEN Test Anxiety Scale

Directions: On your Scantron sheet, fill in the circle that best describes your feeling. Use the scale below to answer each question from A to F, where *A= "Characterizes me perfectly"* and *F= "Does not characterize me at all."*

Characterizes perfectly		**Neutral**		**Does not characterize me me at all**	me
A	B	C	D	E	F

Example: During a test, I keep moving uneasily in my chair.
   A   B   C   D   E   F

*If the student feels this **describes him/her perfectly**, s/he should mark letter "A" on the Scantron sheet.*

1. If I fail a test I am afraid I will be perceived as stupid by my friends.
   A   B   C   D   E   F

2. If I fail a test I am afraid people will consider me worthless.
   A   B   C   D   E   F

3. If I fail a test I am afraid my teachers will look down on me.
   A   B   C   D   E   F

4. If I fail a test I am afraid my teachers will believe I am dumb.
   A   B   C   D   E

5. I am very worried about what my teacher will think or do if I fail his or her test.
   A   B   C   D   E   F

6. I am worried that all my friends will get high scores on the test and only I will get low ones.
   A   B   C   D   E   F

7. I am worried that failure in tests will embarrass me socially.
   A   B   C   D   E   F

8. I am worried that if I fail a test my parents will not like it.
   A   B   C   D   E   F

9. During a test my thoughts are clear and I neatly answer all questions.
   A   B   C   D   E   F

10. During a test I feel I'm in good shape and that I'm organized.
        A   B   C   D   E   F

11. I feel my chances are good to think and perform well on tests.
        A   B   C   D   E   F

12. I usually function well on tests.
        A   B   C   D   E   F

13. I feel I just can't make it on tests.
        A   B   C   D   E   F

14. On a test I feel like my head is empty, as if I have forgotten all I have learned.
        A   B   C   D   E   F

15. During a test it's hard for me to organize what's in my head in an orderly fashion.
        A   B   C   D   E   F

16. I feel it is useless for me to take an examination, I will fail no matter what.
        A   B   C   D   E   F

17. Before a test it is clear to me that I'll fail no matter how well prepared I am.
        A   B   C   D   E   F

18. I am very tense before a test, even if I am well prepared.
        A   B   C   D   E   F

19. While I am sitting in an important test, I feel that my heart pounds strongly.
        A   B   C   D   E   F

20. During a test my whole body is very tense.
        A   B   C   D   E   F

21. I am terribly scared of tests.
        A   B   C   D   E   F

22. During a test I keep moving uneasily in my chair.
        A   B   C   D   E   F

23. I arrive at a test with no serious tension or nervousness.
        A   B   C   D   E   F

Appendix B.

Student Demographic Form

24. Age: 15   16   17   18   19

             A   B   C   D   E

25. Sex:

             A   Male
             B   Female

26. Ethnicity/Race:

             A   Caucasian
             B   Latino or Hispanic
             C   Asian/Pacific Islander
             D   African American
             E   American Indian
             F   Multi-racial

27. Do you receive a free or reduced price lunch?

             A   Yes
             B   No

Directions: The next few questions ask you for basic information about your parent/guardian/caregiver(s). Please answer to the best of your ability.

Parent/Guardian/Caregiver #1

28. Highest level of education achieved:

             A   Grades 0-8
             B   Grades 9-11
             C   High school or GED
             D   Some college/vocational training
             E   College graduate
             F   Graduate/professional degree

29. Current employment:

             A   Yes, full time
             B   Yes, part time
             C   Not working (receiving government assistance)
             D   Not working by choice

Parent/Guardian/Caregiver #2

30. Highest level of education achieved:

             A   Grades 0-8
             B   Grades 9-11

	C	High school or GED
	D	Some college/vocational training
	E	College graduate
	F	Graduate/professional degree

31. Current employment:
    A   Yes, full time
    B   Yes, part time
    C   Not working (receiving government assistance)
    D   Not working by choice

Directions: The next set of questions asks about you and your academic information. Please answer to the best of your ability.

32. Do you receive **special education services** for any of the following?
    A   Learning problem
    B   Vision or hearing problem
    C   Language delay
    D   Emotional/behavioral problem
    E   Other medical problem
    F   ADHD
    G   None

33. Do you receive **gifted education services**?
    A   Yes
    B   No

34. What is your **current semester GPA**? Please rate on a 4.0 scale.
    A   3.50—4.0
    B   3.00—3.49
    C   2.50—2.99
    D   2.00—2.49
    E   1.50—1.99
    F   1.00—1.49

35. What is your **overall/cumulative GPA**? Please rate on a 4.0 scale.
    A   3.50—4.0
    B   3.00—3.49
    C   2.50—2.99
    D   2.00—2.49
    E   1.50—1.99
    F   1.00—1.49

36. Have you previously taken the ACT?
    A   Yes
    B   No

37. Do you plan on having your ACT score reported to colleges?
A   Yes
B   No

38. Do you plan on taking the ACT again?
A   Yes
B   No

39. Have you ever taken an ACT or MME prep class?
A   Yes
B   No

40. What are your future/educational plans? Please indicate **only one** of the following:

A   4-year college
B   2-year community college
C   Vocational school
D   Work force
E   Military or National Guard

Appendix C.

Student Test Anxiety Survey
Directions: The following statements ask you to rate your beliefs, perceptions, and feelings along an Agree-Disagree scale. Please use the scale below to describe how much you agree or disagree with each statement. Fill in **only one** letter for each statement.

Strongly Agree	**Somewhat Agree**	**Neutral**	**Somewhat Disagree**	**Strongly Disagree**
A	B	C	D	E

Example: I am anxious to take the ACT.
*If the student **strongly agrees** with this statement (e.g., feels very anxious to take the ACT), s/he would respond by filling in the letter "**A**" on the Scantron sheet.*

41. I believe my score on the MME is important to my educational/work future.

42. I believe my score on the ACT is important to my educational/work future.

43. I believe my score on the MME (including ACT) is important to my school.

44. I believe my score on the MME (including ACT) is important to my teacher.

45. I am anxious to take the MME.

46. I am anxious to take the ACT.

47. My peers are anxious to take the MME/ACT.

48. My parents are anxious about the MME/ACT.

49. The majority of teachers in my school are anxious about the MME/ACT.

50. The majority of administrators (i.e. principal, assistant principal, superintendent) in my school are anxious about the MME/ACT.

51. People in my community are anxious about the MME/ACT.

52. I believe I am adequately prepared to take the MME/ACT.

53. I believe that my teachers adequately prepared me to take the MME.

54. I am confident about my upcoming performance on MME/ACT.

55. I am **<u>not</u>** nervous to take the MME.

56. I am **<u>not</u>** nervous to take the ACT.

Appendix D.

State-Trait Anxiety Inventory

Directions: A number of statements which people have used to describe themselves are given below. Read each statement and then circle the appropriate number to the right of the statement to indicate **how you feel right now**, that is, **_at this moment._** There are no right or wrong answers. Do not spend too much time on any one statement but give the answer which seems to describe your present feelings best.

Not at all	Somewhat	Moderately so	Very Much
1	2	3	4

57. I feel calm	1	2	3	4
58. I feel secure	1	2	3	4
59. I am tense	1	2	3	4
60. I feel strained	1	2	3	4
61. I feel at ease	1	2	3	4
62. I feel upset	1	2	3	4
63. I am presently worrying over possible misfortunes	1	2	3	4
64. I feel satisfied	1	2	3	4
65. I feel frightened	1	2	3	4
66. I feel comfortable	1	2	3	4
67. I feel self-confident	1	2	3	4
68. I feel nervous	1	2	3	4
69. I am jittery	1	2	3	4
70. I feel indecisive	1	2	3	4
71. I am relaxed	1	2	3	4
72. I feel content	1	2	3	4
73. I am worried	1	2	3	4
74. I feel confused	1	2	3	4
75. I feel steady	1	2	3	4
76. I feel pleasant	1	2	3	4

Directions: A number of statements which people have used to describe themselves are given below. Read each statement and then circle the appropriate number to the right of the statement to indicate how you ___generally feel.___

Not at all	Somewhat	Moderately so		Very Much So
1	2	3		4

77. I feel pleasant	1	2	3	4
78. I feel nervous and restless	1	2	3	4
79. I feel satisfied with myself	1	2	3	4
80. I wish I could be as happy as others seem to be	1	2	3	4
81. I feel like a failure	1	2	3	4
82. I feel rested	1	2	3	4
83. I am "calm, cool, and collected"	1	2	3	4

84. I feel that difficulties are piling up so that I cannot overcome them

	1	2	3	4

85. I worry too much over something that really doesn't matter

	1	2	3	4

86. I am happy	1	2	3	4
87. I have disturbing thoughts	1	2	3	4
88. I lack self-confidence	1	2	3	4
89. I feel secure	1	2	3	4
90. I make decisions easily	1	2	3	4
91. I feel inadequate	1	2	3	4
92. I am content	1	2	3	4

93. Some unimportant thought runs through my mind and bothers me

	1	2	3	4

94. I take disappointments so keenly that I can't put them out of my mind

	1	2	3	4

95. I am a steady person   1   2   3   4

96. I get in a state of tension or turmoil as I think over my recent concerns and interests

	1	2	3	4

Appendix E.

Teacher Survey

School:_____
Primary grades taught:_____
Primary subject taught:_____
Estimated hours of MME prep (please indicate number not range):_____

Please rate the following from 1 to 5 by circling the corresponding number with 1 being strongly disagree, 2 somewhat disagree, 3 neither agree nor disagree, 4 somewhat agree, 5 strongly agree:

1. I believe the MME is important to my students' educational/work future:
    1   2   3   4   5
    strongly   somewhat   neither   somewhat   strongly   disagree
    disagree   agree   agree

2. I believe the ACT is important to my students' educational/work future:
    1   2   3   4   5
    strongly   somewhat   neither   somewhat   strongly   disagree
    disagree   agree   agree

3. I believe the WorkKeys is important to my students' educational/work future:
    1   2   3   4   5
    strongly   somewhat   neither   somewhat   strongly   disagree
    disagree   agree   agree

4. I believe my students' scores on the MME (including ACT and WorkKeys) is important to my school.
    1   2   3   4   5
    strongly   somewhat   neither   somewhat   strongly   disagree
    disagree   agree   agree

5. I believe my students' scores on the MME (including ACT and WorkKeys) is important to my job security.
    1   2   3   4   5
    strongly   somewhat   neither   somewhat   strongly   disagree
    disagree   agree   agree

6. I am anxious for my students' to take the MME:
    1   2   3   4   5
    strongly   somewhat   neither   somewhat   strongly   disagree
    disagree   agree   agree

7. I am anxious for my students' to take the ACT:

1	2	3	4	5	
strongly disagree	somewhat	neither agree	somewhat agree	strongly	disagree

8. I am anxious for my students' to take the WorkKeys:

1	2	3	4	5	
strongly disagree	somewhat	neither agree	somewhat agree	strongly	disagree

9. My peers are anxious for their students to take the MME/ACT/WorkKeys:

1	2	3	4	5	
strongly disagree	somewhat	neither agree	somewhat agree	strongly	disagree

10. The majority of my students' parents are anxious about the MME/ACT/WorkKeys:

1	2	3	4	5	
strongly disagree	somewhat	neither agree	somewhat agree	strongly	disagree

11. The majority of teachers in my school are anxious about the MME/ACT/WorkKeys:

1	2	3	4	5	
strongly disagree	somewhat	neither agree	somewhat agree	strongly	disagree

12. The majority of administrators in my school are anxious about the MME/ACT/WorkKeys:

1	2	3	4	5	
strongly disagree	somewhat	neither agree	somewhat agree	strongly	disagree

13. People in my community are anxious about the MME/ACT/WorkKeys:

1	2	3	4	5	
strongly disagree	somewhat	neither agree	somewhat agree	strongly	disagree

14. I believe I have adequately prepared my students to take the MME/ACT/WorkKeys:

1	2	3	4	5	
strongly disagree	somewhat	neither agree	somewhat agree	strongly	disagree

15. I believe that I had enough time to adequately prepare my students to take the MME:

1	2	3	4	5	
strongly disagree	somewhat	neither agree	somewhat agree	strongly	disagree

16. I believe the MME/ACT/WorkKeys represent a valid assessment of student achievement:

1	2	3	4	5	
strongly disagree	somewhat	neither agree	somewhat agree	strongly	disagree

17. I believe the MME/ACT/WorkKeys represent a valid assessment of teaching effectiveness:

1	2	3	4	5	
strongly disagree	somewhat	neither agree	somewhat agree	strongly	disagree

18. I believe the MME/ACT/WorkKeys represent a valid assessment of school effectiveness:

1	2	3	4	5	
strongly disagree	somewhat	neither agree	somewhat agree	strongly	disagree

19. I have taught my students test anxiety reduction strategies:

1	2	3	4	5	
strongly disagree	somewhat	neither agree	somewhat agree	strongly	disagree

20. I have little time to teaching anything that is not on the MME:

1	2	3	4	5	
strongly disagree	somewhat	neither agree	somewhat agree	strongly	disagree

21. I feel pressure from parents to raise student scores on the MME:

1	2	3	4	5	
strongly disagree	somewhat	neither agree	somewhat agree	strongly	disagree

22. I feel pressure from administrators to raise student scores on the MME:

1	2	3	4	5	
strongly disagree	somewhat	neither agree	somewhat agree	strongly	disagree

23. Administrators in my school feel that student test scores reflect the quality of teaching instruction:

1	2	3	4	5	
strongly disagree	somewhat	neither agree	somewhat agree	strongly	disagree

Appendix F.

Standardized Assessment Directions

My name is _____ and I am from Michigan State University. I am here to learn about student attitudes and feelings about the Michigan Merit Exam. I'm going to be asking you to fill out a brief survey about how you feel towards different parts of the MME and how you feel about tests in general. There are no right or wrong answers. All answers will be strictly confidential and no individual information will be shared will peers, parents, teachers or administrators. As soon as you complete the survey, I will code your responses in a computer and all identifying information will be destroyed.

Please answer the questions to the best of your ability. Again, there are no right or wrong answers. If you have any questions about specific items, please raise your hand and I will do my best to further explain. If you do not know or have an opinion on a specific question, leave it blank.

Thanks for your time.

Appendix G.

Parent Informational Letter

<DATE>

Dear Parent:

Your child is being asked to participate in a research project. Researchers are required to convey that participation is voluntary, to explain risks and benefits of participation, and to empower you to make an informed decision. You should feel free to ask the researchers any questions you may have.

Study Title: HIGH-STAKES ACCOUNTABILITY: EXAMINING STUDENT AND TEACHER ANXIETY WITHIN LARGE SCALE TESTING

Researcher and Title: Nathan von der Embse, Ed.S. Doctoral Candidate and Sara Bolt, Ph.D., Associate Professor

Department and Institution: Department of Counseling, Educational Psychology, and Special Education, Michigan State University

My name is Nathan von der Embse and I am a graduate student at Michigan State University in the School Psychology Program. I am completing my doctoral requirement by examining how students' view the Michigan Merit Exam and whether there is a connection between test performance and test anxiety in select Michigan high schools. All of the eleventh grade students in your child's high school will be asked to participate.

Specifically, your son/daughter will be asked to complete a brief questionnaire, and then their MME results will be reviewed by their high school according to normal procedure then compared to their anxiety results.

Participation should take approximately 15-20 minutes. If you **would not like for your son/daughter to participate**, please complete the attached form. There are no risks involved by participating in this study. Records of this study will be kept confidential, and neither you nor your son/daughter will be identified in any written or verbal reports. Each participant will be coded according to their respective school, teacher and order in which they returned their participation form. All subjects will be assigned a confidential code by the researcher and identifying information will be immediately destroyed.

Scores will be matched according to the coding scheme and any identifying information will be immediately destroyed. Data will only be used for academic and research purposes and no identifying information will be published. All data will be stored on a password protected computer and external hard drive. Any hard copies of data will be stored in a locked room in a locked cabinet only accessible to the researcher.

Please understand that refusal to participate or if you consent and the later withdraw consent or assent from the study will not result in any negative consequences for your child. The researcher will provide the results of the study to each school reported in a summary and no individual data.

By completing the enclosed form you will **have indicated that your son/daughter will be unable to participate in this study**, which will be carried out by myself, under the supervision of Dr. Sara Bolt, Associate Professor in School Psychology, at Michigan State University, East Lansing, Michigan, 517-432-9621. If you have questions about the study, you may direct those to Dr. Bolt at sbolt@msu.edu or myself at 419-303-6781 or vondere1@msu.edu , if you have questions about your rights as a participant, you may contact the Director of MSU's Human Research Protection Programs, Dr. Judy McMillan, at 517-355-2180, FAX 517-432-4503, or e-mail irb@msu.edu, or regular mail at: 207 Olds Hall, MSU, East Lansing, MI 48824. I have enclosed a copy of the opt out/assent forms for you to keep.

Sincerely,

Nathan von der Embse, Ed.S.   Sara E. Bolt, Ph.D.

_____   _____

Appendix H.

RESEARCH OPT OUT FORM

I, (_____), **would _not_ like** (_____) to participate in the research entitled <u>HIGH-STAKES ACCOUNTABILITY: EXAMINING STUDENT AND TEACHER ANXIETY WITHIN LARGE SCALE TESTING</u>, which is being conducted by Nathan von der Embse, Ed.S., (phone number: 419-303-6781). I understand that participation is entirely voluntary and there will be no negative consequences if my child does not participate.

I have read the study description and the following points:

1. **The reason for the research is:** to examine test anxiety differences and test performance within Michigan high schools.

2. **The research procedures are as follows**: The student will complete the FRIEDBEN Test Anxiety Scale (FTAS), the State Trait Anxiety Inventory (STAI), and a brief survey which takes approximately 30 minutes to complete. Surveys will be administered in February one week prior and to taking the Michigan Merit Exam. Forms will be coded and no names will be used. Students' scores on the MME will then be received by the high school and compared with FTAS and survey scores.

3. **The discomforts or stresses that may be faced during this research are:** There are no known discomforts or stress with this research.

4. **The results of this participation will be confidential and will not be released in any individually identifiable form. All forms will be coded by a school administrator and not seen by the researcher.**

Signature of Parent/Guardian: _____Date: _____

Signature of Student Participant _____Date: _____

RETURN THE FORM TO THE RESEARCHER VIA THE SCHOOL.

Appendix I.

<p style="text-align:center">Assent Form</p>

I, (_____), agree to participate in the research entitled <u>HIGH-STAKES ACCOUNTABILITY: EXAMINING STUDENT AND TEACHER ANXIETY WITHIN LARGE SCALE TESTING,</u> which is being conducted by Nathan von der Embse, Ed.S. I understand that this participation is entirely voluntary and there will be no negative consequences if I do not participate. I can withdraw my consent at any time and have the results of the participation removed from any study records.

I have read the study description and the following points are understood:

1. **The reason for the research is:** to examine test anxiety differences and test performance within Michigan high schools.

2. **The research procedures are as follows**: The student will complete the FRIEDBEN Test Anxiety Scale (FTAS), the State Trait Anxiety Inventory (STAI), and a brief survey which takes approximately 20 minutes to complete. Surveys will be administered in February one week prior and to taking the Michigan Merit Exam and eight weeks after taking the exam. Forms will be coded and no names will be used. Students' scores on the MME will then be received by the high school and compared with FTAS and survey scores.

3. **The discomforts or stresses that may be faced during this research are:** There are no known discomforts or stress with this research.

4. **The results of this participation will be confidential and will not be released in any individually identifiable form. All forms will be coded by a school administrator and not seen by the researcher.**

Signature of Student Participant _____Date: _____

Appendix J.

Administrator/Test Coordinator Survey

1. Do your students typically take a "test prep course" or "ACT prep course"?

2. In what setting do the students take the test (classroom, lunchroom, gymnasium)?

3. Does anyone (administrators or teachers) teach test anxiety reduction techniques? If so, what?

4. Do you believe that your staff is anxious for the upcoming test?

5. Do you believe the majority of your students are anxious for the upcoming test?

REFERENCES

REFERENCES

Amrein, A. L., & Berliner, D. C. (2003). The effects of high-stakes testing on student motivation and learning. *Educational Leadership, 60*(5), 32-38.

At a Glance: NCLB and High Schools. Policy Brief (2006). National High School Center.

Balfanz, R., Legters, N., West, T. C., & Weber, L. M. (2007). Are NCLB's measures, incentives, and improvement strategies the right ones for the nation's low-performing high schools? *American Educational Research Journal, 44*(3), 559-593.

Barnard, K. E., Broman-Fulks, J. J., Michael, K. D., Webb, R. M., & Zawilinski, L. L. (2010). The effects of physiological arousal on cognitive and psychomotor performance among individuals with high and low anxiety sensitivity. *Anxiety, Stress & Coping, 24*(2), 201-216.

Beidel, D. C., Turner, M. W., & Trager, K. N. (1994). Test anxiety and childhood anxiety disorders in African American and White school children. *Journal of Anxiety Disorders, 8*(2), 169-179.

Beidel, D. C., & Turner, S. M. (1988). Comorbidity of test anxiety and other anxiety disorders in children. *Journal of Abnormal Child Psychology, 16*(3), 275-287.

Braden, J. P. (2007). Using data from high-stakes testing in program planning and evaluation. *Journal of Applied School Psychology, 23*(2), 129-150.

Bradley, R., McCraty, R., Atkinson, M., Tomasino, D., Daugherty, D., & Arguelles, L. (2010). Emotion self-regulation, psychophysiological coherence, and test anxiety: Results from an experiment using electrophysiological measures. *Applied Psychophysiology and Biofeedback, 35,* 261-283.

Briggs, D., & Domingue, B. (2011). Due diligence and the evaluation of teachers: A review of the value-added analysis underlying the effectiveness rankings of Los Angeles unified school district teachers by the "Los Angeles Times." *National Education Policy Center*. School of Education 249 UCB University of Colorado, Boulder, Colorado 80309.

Bronfenbrenner, U. (1979). *The Ecology of Human Development: Experiments by Nature and Design*. Cambridge, MA: Harvard University Press.

Carter, E. W., Wehby, J., Hughes, C., Johnson, S. M., Plank, D. R., Barton-Arwood, S. M., et al. (2005). Preparing adolescents with high-incidence disabilities for high-stakes testing with strategy instruction. *Preventing School Failure*, *49*, 55-62.

Cizek, G., & Burg, S. (2006). *Addressing Test Anxiety in a High Stakes Environment*. Thousand Oaks, California: Corwin Press.

Clarke, M., Shore, A., Rhoades, K., Abrams, L., Miao, J., Li, J., et al. (2003). Perceived effects of state-mandated testing programs on teaching and learning: Findings from interviews with educators in low-, medium-, and high-stakes states.

Cohen, J. 1977. *Statistical power analysis for the behavioral sciences*. Academic Press, New York, NY.

Connor, M. J. (2003). Pupil stress and standard assessment tasks (SATs): An update. *Emotional and Behavioural Difficulties, 8*(2), 101-107.

Darling-Hammond, L. (2007). Race, inequality and educational accountability: the irony of 'No Child Left Behind'. *Race, Ethnicity & Education, 10*, 245-260.

Diamond, J. B., & Spillane, J. P. (2004). High-stakes accountability in urban elementary schools: Challenging or reproducing inequality? *Teachers College Record, 106*, 1145-1176.

Doyal, G.T. & Forsyth, R.A. (1974). The relationship between student and teacher anxiety levels. *Psychology in the* Schools, *10 (*2), 231-233.

Ediger, M. (2001). *Assessment and high stakes testing*.

Egbochuku, E. & Obodo, B. (2005). Effects of systematic desensitisation (SD) therapy on the reduction of test anxiety among adolescents in Nigerian schools. *Journal of Instructional Psychology, 32,* 298-304.

Ergene, T. (2003). Effective interventions on test anxiety reduction. *School Psychology International, 24*(3), 313-328.

Faber, G. (2010). Enhancing orthographic competencies and reducing domain-specific test anxiety: The systematic use of algorithmic and self-instructional task formats in remedial spelling training. *International Journal of Special Education, 25,* 78-88.

Fielding, C. (2004). Low performance on high-stakes test drives special education referrals: A Texas survey. *Educational Forum, The, 68*(2), 126-132.

Folin, O., Denis, W., & Smillie, W. G. (1914). Some observations on "emotional glycosuria" in man. *Journal of Biological Chemistry, 17*, 519-520.

Friedman, I. A., & Bendas-Jacob, O. (1997). Measuring perceived test anxiety in adolescents: A self-report scale. *Educational and Psychological Measurement, 57*(6), 1035-1046.

Goetz, T., Preckel, F., Zeidner, M., & Schleyer, E. (2008). Big fish in big ponds: A multilevel analysis of test anxiety and achievement in special gifted classes. *Anxiety, Stress & Coping, 21*, 185-198.

Greenberg, M. T., Weissberg, R. P., O'Brein, M. U., Zins, J., Fredericks, L., Resnik, H., et al. (2003). Enhancing school-based prevention and youth development through coordinated social, emotional and academic learning. *American Psychologist*, *58*, 466-474.

Greenberg, P., Sisitsky, T., & Kessler, R. (1999). The economic burden of anxiety disorders in the 1990s. *Journal of Clinical Psychiatry, 60*(6), 427-435.

Gregor, A. (2005). Examination anxiety: Live with it, control it or make it work for you? *School Psychology International, 26*, 617-635.

Hamilton, L., Stecher, B., Marsh, J., McCombs, J.S., Robyn, A., Russell, J., Saftel, S., & Barney, H. (2007). Standards-based accountability under no child left behind: Experiences of teachers and administrators in three states. Santa Monica CA: Rand Corporation. Retrieve from: http://www.rand.org/pubs/monographs/MG589/.

Hamilton, L. S. (2003). Assessment as a policy tool. *Review of Research in Education, 27*, 25-68.

Hanushek, E. A., & Raymond, M. F. (2005). Does school accountability lead to improved student performance? *Journal of Policy Analysis & Management, 24*(2), 297-327.

Hanushek, E. A., & Raymond, M. F. (2004). The effect of school accountability systems on the level and distribution of student achievement. *Journal of the European Economic Association, 2*, 406-415.

Hembree, R. (1988). Correlates, causes, effects, and treatment of test anxiety. *Review of Educational Research, 58*(1), 47-77.

Heubert, J. P., Hauser, R. M., & National Academy of Sciences - National Research Council, W. D. C. (1999). *High stakes: Testing for tracking, promotion, and graduation*.

Hopkins, K.D. & Gullickson, A.R. (1992). Response rates in survey research: A meta-analysis of the effects of monetary gratuities. *Journal of Experimental Education, 61*, 52-62.

Hursh, D. (2005). The growth of high-stakes testing in the USA: Accountability, markets, and the decline in educational equality. *British Educational Research Journal, 31*(5), 605-622.

Jacob, B. (2005). Accountability, incentives and behavior: The impact of high-stakes testing in the Chicago Public Schools. *Journal of Public Economics, 89*(5-6), 761-796.

Jamieson J.P., Mendes W.B., Blackstock E., & Schmader T. (2010).Turning the knots in your stomach into bows: Reappraising arousal improves performance on the GRE. *Journal of Experimental Social Psychology, 46*(1), 208-212.

Joesting, J. (1975). Test-retest reliabilities of State-Trait Anxiety Inventory in an academic setting. *Psychological Reports, 37*(1), 270.

Katsiyannis, A., Zhang, D., Ryan, J. B., & Jones, J. (2007). High-stakes testing and students with disabilities. *Journal of Disability Policy Studies, 18*, 160-167.

Kessler, R., Chiu, W., Demler, O., Merikangas, K., & Walters, E. (2005). Prevalence, severity, and comorbidity of 12 month DSM-IV disorders in the national comorbidity survey replication. *Archives of General Psychiatry, 62*(6), 617-626.

King, N. J., & Ollendick, T. H. (1989). Children's anxiety and phobic disorders in school settings: Classification, assessment, and intervention issues. *Review of Educational Research, 59*(4), 431-470.

Koretz, D. M., & Hamilton, L. S. (2006). Testing for accountability in K-12. In R. L. Brennen (Ed.), *Educational Measurement* (4 ed., pp. 531-542). Westport, CT: American Counsil on Education/Praeger.

Lang, J. & Lang, J. (2010). Priming competence diminishes the link between cognitive test anxiety and test performance: Implications for the interpretation of test scores. *Psychological Science, 21,* 811-819.

Larson, H., Ramahi, M., Conn, S., Estes, L., & Gibellini, A. (2010). Reducing test anxiety among third grade students through the implementation of relaxation techniques. *Journal of School Counseling, 8,* 1-19.

Legrand, L., McGue, M., & Iacono, W. G. (1999). A twin study of state and trait anxiety in childhood and adolescence. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 40*(6), 953-958.

Liebert, R. M., & Morris, L. W. (1967). Cognitive and emotional components of test anxiety: a distinction and some initial data. *Psychological Reports, 20*(3), 975-978.

Linn, R. L. (2000). Assessments and accountability. *Educational Researcher, 29*(2), 4-17.

Lowe, P. A., & Lee, S. W. (2008). Factor structure of the Test Anxiety Inventory for Children and Adolescents (TAICA): Scores across gender among students in elementary and secondary school settings. *Journal of Psychoeducational Assessment, 26*(3), 231-246.

Lowe, P. A., Lee, S. W., Witteborg, K. M., Prichard, K. W., Luhr, M. E., Cullinan, C. M., et al. (2008). The Test Anxiety Inventory for Children and Adolescents (TAICA): Examination of the psychometric properties of a new multidimensional measure of test anxiety among elementary and secondary school students. *Journal of Psychoeducational Assessment, 26*(3), 215-230.

Mandler, G., & Sarason, S. B. (1952). A study of anxiety and learning. *Journal Of Abnormal Psychology, 47*(2), 166-173.

Marsh, H. W. (1987). The big-fish-little-pond effect on academic self-concept. *Journal of Educational Psychology, 79*(3), 280-295.

Marsh, H. W., Trautwein, U., Ludtke, O., & Koller, O. (2008). Social comparison and big-fish-little-pond effects on self-concept and other self-belief constructs: Role of generalized and specific others. *Journal of Educational Psychology, 100*(3), 510-524.

Marteau, T.M., & Bekker, H. (1992). The development of a six-item short-form of the state scale of the Spielberger State-Trait Anxiety Inventory (STAI). *British Journal of Clinical Psychology, 3,* 301-306.

Mathis, William J. (2011) "Race to the Top: An example of belief-dependent reality. *Democracy and Education*, *19*(2), 1-7.

Mayer, J. D., Roberts, R. D., & Barsade, S. G. (2008) Human abilities: Emotional intelligence. *Annual Review of Psychology, 59¸* 507-536.

McDonald, A. S. (2001). The prevalence and effects of test anxiety in school children. *Educational Psychology, 21*, 89-101.

Morris, L. W., & Liebert, R. M. (1970). Relationship of cognitive and emotional components of test anxiety to physiological arousal and academic performance. *Journal Of Consulting And Clinical Psychology, 35*(3), 332-337.

Mullis, I. V. S., Martin, M. O., Beaton, A. E., Gonzalez, E. J., Kelly, D. L., Smith, T. A., et al. (1998). *Mathematics and science achievement in the final year of secondary school: IEA's third international mathematics and science study (TIMSS).*

Neal, D., Schanzenbach, D. W., & National Bureau of Economic Research, C. M. A. (2007). Left behind By design: Proficiency counts and test-based accountability. NBER Working Paper No. 13293.

Nicholson-Crotty, S., & Staley, T. (2012). Competitive federalism and Race to the Top Application decisions in the American states. *Educational Policy, 26*(1), 160-184.

No Child Left Behind Act of 2001, P.L. 107-110 C.F.R.

Onosko, Joe (2011) "Race to the Top leaves children and future citizens behind: The devastating effects of centralization, standardization, and high stakes accountability. *Democracy and Education*, *19*(2), 1-11.

Pallant, J. (2007). *SPSS Survival Manual: A step by step guide to data analysis using SPSS for Windows, third edition.* Berkshire, England: Open University Press.

Pedulla, J. J., Abrams, L. M., Madaus, G. F., Russell, M. K., Ramos, M. A., Miao, J., et al. (2003). Perceived effects of state-mandated testing programs on teaching and learning: Findings from a national survey of teachers.

Peleg, O. (2009). Test anxiety, academic achievement, and self-esteem among Arab adolescents with and without learning disabilities. *Learning Disability Quarterly, 32*, 11-20.

Putwain, D. (2007). Test anxiety in UK schoolchildren: Prevalence and demographic patterns. *British Journal of Educational Psychology, 77*, 579-593.

Putwain, D. (2008a). Deconstructing test anxiety. *Emotional & Behavioural Difficulties, 13*, 141-155.

Putwain, D. (2008b). Do examinations stakes moderate the test anxiety-examination performance relationship? *Educational Psychology, 28*, 109-118.

Putwain, D. (2008c). Test anxiety and GCSE performance: The effect of gender and socio-economic background. *Educational Psychology in Practice, 24*, 319-334.

Raffety, B. D., Smith, R. E., & Ptacek, J. T. (1997). Facilitating and debilitating trait anxiety, situational anxiety, and coping with an anticipated stressor: A process analysis. *Journal of Personality & Social Psychology, 72*, 892-906.

Resnick, D. (1980). Minimum competency testing historically considered. *Review of Research 77*, 579-593.

Rosairio, P., Naez, J. C., Salgado, A., Gonzalez-Pienda, J. A., Valle, A., Joly, C., et al. (2008). Test anxiety: associations with personal and family variables. *Psicothema, 20*(4), 563-570.

Rothstein, J. (2010). Teacher Quality in Educational Production: Tracking, Decay, and Student Achievement*. *Quarterly Journal of Economics, 125*(1), 175-214.

Ryan, K. E., Ryan, A. M., Arbuthnot, K., & Samuels, M. (2007). Students' motivation for standardized math exams. *Educational Researcher, 36*(1), 5-13.

Santos, F., & Otterman, S. (2012, February 24). City teacher data reports are released. *The New York Times*. Retrieved from http://www.nytimes.com.

Sarason, S. B., Davidson, K., Lighthall, F., & Waite, R. (1958). A test anxiety scale for children. *Child Development, 29*(1), 105-113.

Sarason, S. B., Hill, K. T., & Zimbardo, P. (1964). A longitudinal study of the relation of test anxiety to performance on intelligence and achievement tests. *Monographs of the Society for Research in Child Development, 29*(7), 1-51.

Sarason, S. B., & Mandler, G. (1952). Some correlates of test anxiety. *Journal Of Abnormal Psychology, 47*(4), 810-817.

Schutz, P. A., & Davis, H. A. (2000). Emotions and self-regulation during test taking. *Educational Psychologist, 35*, 243-256.

Schutz, P. A., Distefano, C., Benson, J., & Davis, H. A. (2004). The emotional regulation during test-taking scale. *Anxiety, Stress & Coping, 17*, 253-269.

Segool, N. K. (2009). *Test anxiety associated with high-stakes testing among elementary school children: Prevalence, predictors, and relationship to student performance.* Michigan State University. *ProQuest Dissertations and Theses,* http://ezproxy.msu.edu/login?url=http://search.proquest.com/docview/304940211?accountid=12598

Sena, J. D. W., Lowe, P. A., & Lee, S. W. (2007). Significant predictors of test anxiety among students with and without learning disabilities. *Journal of Learning Disabilities, 40*, 360-376.

Smith, M. (1991). Put to the Test: The Effects of External Testing on Teachers. *Educational Researcher, 20*, 8-11.

Song, J., & Felch, J. (2011, May 7). Times updates and expands value-added ratings for Los Angeles elementary school teachers. *The Los Angeles Times*. Retrieved from http://www.latimes.com.

Spielberger, C. D. (1966). Theory and research in anxiety. In C. D. Spielberger (Ed.), *Anxiety and Behavior* (pp. 3-20). New York: Academic Press.

Spielberger, C. D., & Vagg, P. R. (1984). Psychometric properties of the STAI: A reply to Ramanaiah, Franzen, and Schill. *Journal of Personality Assessment, 48*, 95.

Stanton, H. (1974). The relationship between teachers' anxiety levels and the test anxiety level of their students. *Psychology in the* Schools, *11 (*3), 360-363.

Stöber, J., & Pekrun, R. (2004). Advances in test anxiety research. *Anxiety, Stress & Coping*, 205-211.

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (5th ed.). Boston: Allyn and Bacon.

Thibodeau, M. A., Gómez-Pérez, L., & Asmundson, G. J. G. (2012). Objective and perceived arousal during performance of tasks with elements of social threat: The influence of anxiety sensitivity. *Journal of Behavior Therapy and Experimental Psychiatry, 43*(3), 967-974.

Triplett, C. F., & Barksdale, M. A. (2005). Third through sixth graders' perceptions of high-stakes testing. *Journal of Literacy Research, 37*(2), 237-260.

Turner, B. G., Beidel, D. C., Hughes, S., & Turner, M. W. (1993). Test anxiety in African American school children. *School Psychology Quarterly, 8*(2), 140-152.

von der Embse, N., Barterian, J., & Segool, N. (in press). Test anxiety interventions for children and adolescents: A systematic review of treatment studies from 2000-2010. *Psychology in the Schools*.

Walsh, K., & Jacobs, S. (2009). Race to the Top: Colorado may be used to high altitudes but can it compete in Race to the Top? *National Council on Teacher Quality*. 1420 New York Avenue NW Suite 800, Washington, DC 20005.

Weems, C., Scott, B., Taylor, L., Cannon, M., Romano, D., Perry, A., & Triplett, V. (2010). Test anxiety prevention and intervention programs in schools: Program development and rationale. *School Mental Health, 2,* 62-71.

Williams, J. E. (1996). Gender-related worry and emotionality test anxiety for high-achieving students. *Psychology in the Schools, 33*(2), 159-162.

Yahav, R. & Cohen, M. (2008). Evaluation of a cognitive-behavioral intervention for adolescents. *International Journal of Stress Management, 15,* 173-188.

Yerkes, R. M., & Dodson, J. D. (1908). The relation of strength stimulus to rapidity of habit-formation. *Journal of Comparative and Neuropsychological Psychology, 18*, 459-482.

Ysseldyke, J., Nelson, J. R., Christenson, S., Johnson, D. R., Dennison, A., Triezenberg, H., et al. (2004). What we know and need to know about the consequences of high-stakes testing for students with disabilities. *Exceptional Children, 71*, 75-95.

Zeidner, M. (1990). Does test anxiety bias scholastic aptitude test performance by gender and sociocultural group? *Journal of Personality Assessment, 55*, 145.

Zeidner, M. (1998). *Test Anxiety, The State of the Art*. New York: Plenum Press.