

2011

# California High School Exit Exam for students with disabilities : the impact of setting, anxiety, and stereotype threat on stuent's math performance

Tracy M. Fields

Follow this and additional works at: <https://repository.usfca.edu/diss>

 Part of the [Education Commons](#)

---

## Recommended Citation

Fields, Tracy M., "California High School Exit Exam for students with disabilities : the impact of setting, anxiety, and stereotype threat on stuent's math performance" (2011). *Doctoral Dissertations*. 282.  
<https://repository.usfca.edu/diss/282>

This Dissertation is brought to you for free and open access by the Theses, Dissertations, Capstones and Projects at USF Scholarship: a digital repository @ Gleeson Library | Geschke Center. It has been accepted for inclusion in Doctoral Dissertations by an authorized administrator of USF Scholarship: a digital repository @ Gleeson Library | Geschke Center. For more information, please contact [repository@usfca.edu](mailto:repository@usfca.edu).

The University of San Francisco

CALIFORNIA HIGH SCHOOL EXIT EXAM FOR STUDENTS WITH  
DISABILITIES: THE IMPACT OF SETTING, ANXIETY,  
AND STEREOTYPE THREAT ON STUDENTS'  
MATH PERFORMANCE

A Dissertation Presented  
to  
The Faculty of the School of Education  
Learning and Instruction Department

In Partial Fulfillment  
of the Requirements for the Degree  
Doctor of Education

by  
Tracy M. Fields  
San Francisco  
May 2011

THE UNIVERSITY OF SAN FRANCISCO

Dissertation Abstract

California High School Exit Exam for Students with Disabilities: The Impact of Setting, Anxiety, and Stereotype Threat on Students' Math Performance

In many states across the U.S. there is a large percent of students with disabilities who exit high school without a diploma. One of the barriers to receiving a high school diploma for students with disabilities is the passage of an exit exam. Anxiety and stereotype threat have been seen as mechanisms or factors in explaining low performance on standardized tests.

Research with regards to students with disabilities academic achievement based on instructional setting has been mixed. Thus, the purpose of this study was to examine the impact of setting, anxiety, and stereotype threat on the math scores of the California High School Exit Exam for students with disabilities who receive math instruction in a mainstream or self-contained setting.

A causal comparative research design was used to compare two settings, mainstream and self-contained, with respect to the three variables (anxiety, stereotype threat, and CAHSEE). Students with disabilities who took the CAHSEE for the first time in the spring administration of the test were utilized. Since student participants were in intact groups, no random assignment of participants was conducted.

After the administration of the mathematic section of the CAHSEE students with disabilities were asked to complete a questionnaire to measure test anxiety, and to indirectly measure stereotype threat. A one-way analysis-of-covariance (ANCOVA) was conducted to answer each research question. A principal component analysis (PCA) was done to transform a set of correlated variables into one variable described as prior math ability which was used as a covariate.

A major finding of the study indicates that anxiety predicts CAHSEE scores whereas stereotype threat does not. Although anxiety was found to be highly correlated with CAHSEE math performance, anxiety and stereotype threat do not significantly impact the scores on the math section on the CAHSEE between students with disabilities in mainstream or self-contained settings for math instruction. Another major finding suggests that the setting in which students with disabilities received math instruction (mainstream or self-contained) has an impact on test scores on the math section of the CAHSEE. Implications for future research and practice are discussed.

This dissertation, written under the direction of the candidate's dissertation committee and approved by the members of the committee, has been presented to and accepted by the Faculty of the School of Education in partial fulfillment of the requirements for the degree of Doctor of Education. The content and research methodologies presented in the work represent the work of the candidate alone.

Tracy M. Fields	5/9/2011
_____	_____
Candidate	Date
Dissertation Committee	
Dr. Yvonne Bui	5/9/2011
_____	_____
Chairperson	
Dr. Robert Burns	5/9/2011
_____	_____
Dr. Patricia Mitchell	5/9/2011
_____	_____

## DEDICATION

I dedicate this dissertation to the many friends and family members who have supported me through this journey. To my husband Fitzgerald, my daughter Taylor, and my sons Barry and Austin, We Did It! Thank you for all the support you have given me. We did this together. Know that I love you with all my heart. To all my students, don't ever give up on your goals. Sometimes it will take longer than expected, but if you can believe, conceive, and persevere, then you can achieve.

## ACKNOWLEDGEMENTS

I would like to take this opportunity to acknowledge those individuals who have been supportive in my dissertation efforts. I wish to thank the members of my dissertation committee, especially my chair Dr. Yvonne Bui, for her guidance, support, patience, and confidence in me. To Dr. Robert Burns and Dr. Patricia Mitchell thank you for your support and dedication throughout this process. A special thank you to Dr. Susan Evans and Dr. Kimberly Mayfield-Lynch for without you, this journey would not have begun; and to Shannon Halkyard, thank you for your technical guidance and support.

I want to extend my appreciation to the students and staff members of Castro Valley, Dublin, and New Haven Unified School Districts as well as Capuchino High School. Without your support, this would not be possible. To Dr. Nicholas and Angela Baham, thank you for your guidance and prayers. I will always be grateful. To my sorors of Alpha Kappa Alpha Sorority, Inc. Rho Upsilon Omega chapter, thank you for allowing me to take time off to complete this dissertation and for your thoughtful wishes and encouragement. To the members of Jack and Jill of America Tri-Valley chapter, thank you for allowing me the time need to finish this race.

To all my friends who have helped me in some way whether known or unknown, just know that I appreciate you. To the Allen Temple Baptist Church family, thank you for your prayers and encouragement. Although the times got rough, you stayed with me, prayed with me, and even cried with me. You continue to lift me up. Again I say thank you.

Most importantly I want to say thank you to my family. To my mother, thank you for always being in my corner. To my father who knew I could achieve whatever I set my mind to, thank you for your support. To my sister Sylvia and my brothers Martin and Russell, you too can do it. Never give up, keep on striving, and don't stop praying. I love you guys. To my husband thank you for sticking with me through thick and thin. To my children, you are my light. To family who are not here with me to share in this triumph, thank you for your love and support. I know you are rejoicing from above. It takes a village to raise a child, thank you to my village. I would never have made it without you. To God be the glory, Forever!

## TABLE OF CONTENTS

	Page
TABLE OF CONTENTS	viii
LIST OF TABLES	xi
LIST OF FIGURES	xii
CHAPTER ONE: STATEMENT OF THE PROBLEM	1
Purpose of the Study	5
Significance of the Study	6
Background and Need	7
High standards and high stakes.	8
Gaps in passage and graduation rates.	12
Math ability and setting.	17
Theoretical Rationale	18
Research Questions	22
Definition of Key Terms	23
CHAPTER TWO: LITERATURE REVIEW	26
Mainstreaming and Self-Contained Settings	26
High Stakes Testing for Students with Disabilities	40
Stereotype Threat	61
Summary of the Review of the Literature	80
CHAPTER THREE: METHODOLOGY	82
Research Design	83
Setting	84
Sample	86
Protection of Human Subjects	88
Instrumentation	89
Math section of CAHSEE	90
Anxiety instrument	91

Stereotype threat instrument.	91
Pilot Study	92
Validity Panel	93
Procedures	95
Data collection.	96
Questionnaire administration.	97
School record review.	98
Data analysis.	99
Summary	100
CHAPTER FOUR: RESULTS	102
Descriptive Statistics	102
Analysis Related to Research Question 1	107
Analysis Related to Research Question 2	108
Analysis Related to Research Question 3	109
Summary	110
CHAPTER FIVE: SUMMARY, LIMITATIONS, DISCUSSION, CONCLUSIONS, AND IMPLICATIONS	112
Summary of Findings	114
Limitations	115
Discussion of Findings	117
Conclusions	123
Implications	126
For research.	126
For practice.	127
Summary	128
REFERENCES	130
APPENDIXES	138
Appendix A- USF IRB Approval	139
Appendix B- Student Questionnaire Coversheet	140
Appendix C- Student Anxiety Questionnaire	141

Appendix D- Student Stereotype Threat Questionnaire	142
Appendix E- Validity Panel Questionnaire	143
Appendix F- Validity Panel Demographics	144
Appendix G- Validity Panel Packet	145
Appendix H- Validity Panel Comments	148
Appendix I- School District Consent	151
Appendix J- Parental Consent Packet	155
Appendix K- Subjects' Bill of Rights	158
Appendix L- Student Data Coding Sheet	159

## LIST OF TABLES

	Page
Table 1: Summary Demographics of Participating Schools	85
Table 2: Frequency Table of Demographics for Participants	87
Table 3: Means, Standard Deviations (SD), and Sample Sizes (n) of California High School Exit Exam (CAHSEE) Math Scores, Stereotype Threat (STT), Anxiety, Overall 9th Grade Point Average (GPA), and Prior Path Ability (PMA) Components Scores for Mainstream and Self-Contained Groups	103
Table 4: Correlations among Anxiety, Stereotype Threat (STT), Prior Path Ability (PMA) Components Scores, Overall 9th Grade Point Average (GPA), and California High School Exit Exam (CAHSEE) Math Scores (N = 64)	105
Table 5: Multiple Linear Regression of Anxiety, Stereotype Threat (STT) and Prior Math Ability (PMA)	106
Table 6: Analysis of Covariance Summary for California High School Exit Exam by Setting	108
Table 7: Analysis of Covariance Summary for Stereotype Threat by Setting	109
Table 8: Analysis of Covariance Summary for Anxiety by Setting	110
Table 9: Summary of Analysis of Covariance Mean and Adjusted Mean for California High School Exit Exam (CAHSEE), Anxiety, and Stereotype Threat (STT) by Setting	111

## LIST OF FIGURES

	Page
Figure 1: The impact of setting, stereotype threat, and anxiety on CAHSEE scores	84

## **CHAPTER ONE**

### **STATEMENT OF THE PROBLEM**

In many states across the U.S. there is a large percent of students with disabilities who exit high school without a diploma. According to the national average in 2006, 70% of students without disabilities graduated with a high school diploma. Within that same year, 57% of students with disabilities graduated with a diploma nationwide (Planty et al., 2008). One of the barriers to receiving a high school diploma for students with disabilities is the passage of an exit examination requirement. Currently 26 states have exit exam policies for graduation (Center on Education Policy, 2007). By 2012, it is estimated that 70 percent of students nationwide will take exit exams as a mandatory graduation requirement for a high school diploma (Center on Education Policy, 2004, 2005). This current push for exit exam requirements, which received support from various federal actions and legislation, is part of a movement towards standards-based reform.

Exit exams are a vital part of the standards-based reform movement. They are used to show that students are making progress towards grade level standards as a guarantee to the public that a student with a diploma has acquired the skills and knowledge needed for the workplace, college, and other aspects of life (Center on Education Policy, 2002). The exit exam is also an accountability tool used to ensure the high quality of public education in the U. S. Under the federal No Child Left Behind (NCLB) Act of 2001, all students are to achieve academic proficiency in reading/language arts and math by 2013. Schools, districts, and states show their compliance with NCLB by making Adequate Yearly Progress (AYP) measured by

students meeting their state's academic standards as demonstrated through state assessments (California Department of Education, 2009). For example, California's AYP is derived from assessment results on the Standardized Testing and Reporting (STAR) Program, and the California High School Exit Examination (CAHSEE). The student's first time CAHSEE score is important to a school's accountability report. A student's first attempt, usually in the spring of their sophomore year in high school, is the measure used to calculate the school's AYP and what California includes in its accounting report to the federal government.

Students most affected by the requirement of an exit exam for graduation with a diploma are those who fail their first attempt. These students are usually given remedial coursework in place of general electives or provided tutoring programs to assist them in successful passage on future attempts. Depending on the state, up to 70% of students fail the mathematics portion of the test on their first attempt and up to 40% fail the English/language arts portion (Center on Education Policy, 2007). The percentages of students who fail on their first undertaking include a significant number of minority students and students with disabilities (Center on Education Policy, 2007; Horn, 2003). For example, in California the pass rate on the exit exam for 10<sup>th</sup> grade students who took the exam for the first time in 2009 was 42% for English learners, 72% for Hispanic or Latino, 70% for African-Americans, and 37% for students with disabilities compared to 80% for all students (California Department of Education, 2009).

Consequences have been attached to students' failure on the high school exit exams. As such, high school exit exams are considered "high stakes" tests for both the school district and the student. In California, exit exams have an effect on a school meeting AYP and the Academic Performance Index (API) which is an additional state

accountability requirement. For example, the school could be labeled in need of performance improvement or taken over by state officials if it does not meet API and AYP (California Department of Education, 2009). For students, the exit exam affects their ability to graduate with a high school diploma. Not having a high school diploma can directly affect an individual's future economic self-sufficiency and well-being as an adult (Johnson & Thurlow, 2005; Manset & Washburn, 2000). As stated in O'Neill (2001), individuals without a high school diploma or General Education Diploma (GED) earns approximately 19% less per hour than an individual with a diploma. Thus, the pressure to pass the high stakes exit exam can be enormous, especially for minority students, English learners, economically disadvantaged students, and students with disabilities who historically have done poorer on standardized tests (Center on Education Policy, 2005).

Research published about high school exit exams as a graduation requirement has been diverse in scope. Some of the literature has focused on the accommodations needed for students with disabilities to be successful on the exam (Chapman v. California Department of Education, 2002; Johnson, Thurlow, & Stout, 2007). Other work has focused on questioning the validity of the exams or its effects on student achievement, persistence in high school, graduation rates, and standardized tests (Brookhart v. Illinois State Board of Education, 1983; Callet, 2005; Debra P. v. Turlington, 1981; Reardon, Atteberry, Arshan, & Kurlaender, 2009). For instance, the validity of the test scores may be compromised because the pictures and charts used in the reading sections may be distracting and unfamiliar to some students. In addition, the ability to assess certain standards the test purports to measure may be limited due to the inflexibility in the format of the multiple choice response section. The inability to measure the standards would

have an impact on the scores and may compromise the validity of the test (Callet, 2005). Although much has been written about high school exit exams as a mandatory graduation requirement for all students, little has been written to address the issue of why students with disabilities continue to fail the first time at alarming rates. Therefore, this study addressed the gap in the literature regarding an important population in the public schools, students with disabilities.

High school exit exams are high stakes testing situations because of the consequences for students who do not pass the exam. Students who do not pass the high school exit exam do not graduate with a high school diploma. All states that have exit exam policies report gaps in pass rates among various groups of students (Center on Education Policy, 2007) with pass rates on the high school exit exam being lower for students with disabilities. One possible reason is that students with disabilities may experience an increased anxiety associated with stereotype threat, thereby scoring lower on the exam. Research suggests that stereotype threat causes increased anxiety about confirming a negative stereotype about a group to which the student belongs (Steele, 1997; Steele, Spencer, & Aronson, 2002) thus directly interfering with the students' test performance.

Special educators have long known about problems of stereotyping, leading to the use of inclusive settings versus self contained settings. Inclusive settings, called mainstream settings, are thought to improve student outcomes academically and socially compared to self-contained settings (Holloway, 2001; Rea, McLaughlin, & Walther-Thomas, 2002). Inclusive settings are looked upon as a means of reducing anxiety and increasing a students' self-concept. Therefore, it may be possible that placement of a student with disabilities in an inclusive settings insulates the student against anxiety and

stereotype threat. Unfortunately, there is little research on this possibility. Therefore, the purpose of this study was to examine the hypothesis that inclusive settings help reduce anxiety and stereotype threat, thereby increasing performance in a high stakes testing situation.

### **Purpose of the Study**

The purpose of the study was to investigate students' anxiety associated with stereotype threat in relation to California High School Exit Exam (CAHSEE) scores in mathematics for student with disabilities in mainstream or self-contained classroom setting for math instruction. This study will assist educators in understanding the role anxiety and stereotype threat, in regards to setting, affect student performance prior to a student's first attempt on the CAHSEE with the intent of increasing student achievement scores on the exam. This study expands the body of research related to exit exams and focus not on the exit exam itself, but differences in performance between students with disabilities who were in different academic settings for math instruction, and the threat of stereotypes on the success or failure on the first attempt of the mathematics section on the CAHSEE.

Students who do not receive a passing score on the CAHSEE are in jeopardy of not receiving their high school diploma. A student who does not pass the first attempt of the exit exam may be placed in remedial classes instead of an elective in preparation for the next administration of the test. Without a diploma, students limit their access to post-secondary education, the military, or other higher level career options (Albrecht & Joles, 2003). In addition, districts face state and federal accountability requirements to meet Adequate Yearly Progress (AYP) and the state's Academic Performance Index (API). The results on the CAHSEE are used for measuring a school's progress towards meeting

these accountability requirements. As such, the information will assist school districts in addressing issues related to success on the high school exit exam.

### **Significance of the Study**

Exit exams have a tremendous impact on the post high school achievement of students with disabilities because high school exit exams determine the likelihood that a student will receive a high school diploma in California. Currently, a gap continues to exist in pass rates on the CAHSEE for students with disabilities and general education students who take the test for the first time (California Department of Education, 2008c). This study explored school setting, anxiety, stereotype threat and results on the CAHSEE for students with disabilities to assist in increasing pass rates. Increased pass rates of students with disabilities will aid in a school and its district's accountability reports, relieving the pressure felt from the state and the community when students are not making progress towards state goals.

For students with disabilities, the possession of a high school diploma will enable them to enter the work force at a high wage and continue their education on the job or obtain additional education in a supported college program (Albrecht & Joles, 2003). As a result of continuous failure on the CAHSEE, students may drop out of high school. Students who drop out of high school are less likely to be in the work force than those who possess a high school diploma or higher and underemployed in the work force earning low wages (Laird, DeBell, Kienzl, & Chapman, 2007). Lower earnings among dropouts alone could cost the United States as much as \$158 billion in lost earnings and \$36 billion in lost state and federal income taxes for each class of 18-year-olds. This amount is estimated to account for 1.6% of the nation's gross national gross domestic product (Rouse, 2005). With regards to health, the differences between high school

graduates and dropouts are evident. Dropouts over the age of 24 tend to have poorer health than adults who are not dropouts, regardless of their income status (U.S. Department of Education, 2004). Muennig (2005) reported that high school dropouts have an average of \$35,000 in annual health care costs and live an average of nine fewer years than graduates. Without a diploma, students' options for making a living is limited, creating a permanent lower class of uneducated, unemployed citizens and residents (Garcia, 2003).

To investigate the issue of students not passing the CAHSEE and thus not receiving a diploma, this study attempted to use the stereotype of receiving special education services and Steele and Aronson's (1995) stereotype threat theory in the investigation of success or failure rates on the CAHSEE for first-time 10<sup>th</sup> grade examinees with disabilities in northern California school districts. This study is significant in that it addresses the issue of setting, anxiety, and stereotype threat as having an impact on the CAHSEE scores. The study enlightens educational policy related to students with disabilities. Moreover, this study looks to support and enhance previous research related to standardized high stakes testing and stereotype threat for students with disabilities. Since the first time examination results are the ones used in the accountability reports and given the possible negative consequences associated with failure on high school graduation exams, factors that impact CAHSEE scores needed to be investigated for those who are affected by the results.

### **Background and Need**

In today's educational system, higher standards are being set for students to receive a high school diploma. In many states, students with disabilities who want to

receive a diploma when exiting high school will need to take and pass an exit exam. If a student does not pass the exit exam consequences exist for both the student and the school district. For certain groups, the pass rates on the exit exams are low. Although gaps exist in pass rates among students, students who have disabilities are the most at-risk for not receiving a diploma.

Research has been conducted to address the gap in achievement and to identify factors predictive of success on the exit exam, but little has been done to address the gap affecting students with disabilities. This study addressed this gap in the literature. The following sections present information surrounding high school exit exam issues specific to the United States and California. Organized into three subsections, High Standards and High Stakes, Gap in Passage and Graduation Rates, and Math Ability and Setting, the subsections provided a rationale for the investigation of exit exam passage rates for students with disabilities.

**High standards and high stakes.** To date, approximately 22 states have implemented the passage of a mandatory standardized test to receive a diploma with other states planning to enforce this requirement over the next several years (Center on Education Policy, 2007). The exams are considered “high-stakes” tests due to the consequences attached for individuals unable to receive the minimum score for passage (Manset & Washburn, 2000). Consequences include the denial of a high school diploma, regardless of whether the student has the grades and completed all course requirements, the possible embarrassment caused by having to retake the test, and the time and effort spent in remediation activities.

The impetus of high school exit exams began from public criticism of the educational system. State policymakers adopted the use of exit exams in response to

increased concerns by business leaders, college faculty, and others that the quality of public education had diminished and that young people were graduating with poor academic skills (Center on Education Policy, 2002). For example in 2002, in a poll conducted by Public Agenda, 7 out of 10 employers and professors perceived that young people (post high school age) have fair or poor skills in grammar, spelling, and writing, and approximately 6 of 10 high school graduates have fair or poor skills in basic mathematics (Center on Education Policy, 2002).

The purpose for exit exams has changed over time. Previously states reported the purpose for exit exams as preparing students for life after high school, readiness for post-secondary education, or as a way to determine mastery of the state curriculum (e.g., standards, curriculum frameworks). Today the majority of states with exit exams view them as a way of determining mastery of state curriculum (e.g. standards, curriculum frameworks), increasing alignment of local curriculum and instruction with state education standards, and meeting the high school assessment requirements of the No Child Left Behind Act (NCLB) of 2001 (Center on Education Policy, 2007).

With the reauthorization of Individuals with Disabilities Education Act (IDEA) in 2004, and its alignment with the No Child Left Behind (NCLB) Act of 2001, students with disabilities are held to the same accountability system as students without disabilities. Federal laws allow students with disabilities and English language learners to use appropriate test accommodations as needed and documented. In 2007, California's Board of Education recommended that all students take and pass the California High School Exit Exam (CAHSEE) and that California not develop alternative assessments (Samuels, 2007). The Board of Education stated that certificates of attendance or achievement would be given to students with disabilities who were unable to pass the test

or decided not to take it. Most states, as well as California, with exit exam policies require all students to satisfactorily complete their coursework and pass the exam in order to receive a diploma.

Students who complete the coursework but fail the exam are in jeopardy of not receiving a high school diploma. If students do not pass the exit exam after multiple tries, they either receive a “certificate of completion” or “certificate of attendance” (Center on Education Policy, 2002). Without a high school diploma, outcomes for adult life are less favorable as compared to those who do receive a diploma. Failing to complete high school with a diploma, due to a graduation test or other reasons, tend to result in problems associated with employment, earnings, family stability, and health (Hauser, 1997). Students who graduated from high schools where states had instituted exit exam policies tended to obtain higher paying jobs and were able to sustain their pay advantage over a five year period (Bishop, Mane, & Bishop, 2001).

The subject areas assessed on the exit exams vary slightly from state to state. Some include science or social science, but the most common subjects tested are reading, writing, and mathematics. Exit exam items include multiple-choice questions and some form of essay writing. All states with exit exams provide multiple opportunities, from two up to eight tries, to retake the exam. The retake tests are parallel versions of the test with slightly different questions (Center on Education Policy, 2002). By their senior year, about 90% of high school students ultimately pass the test, but a large percent of students fail on their first attempt. An unknown number of students drop out before ever taking the test due to differences in state reporting. Repeat test takers may drive up the cost associated with the implementation of an exit exam policy (Center on Education Policy, 2007).

The cost associated with exit exams poses consequences for the school district as well as the student. An estimated 96% of the cost associated with the implementation of exit exam policies is paid for by the local school district. There are costs associated with remedial services for students who fail the exams and the implementation of prevention programs to prevent failure on the exams. Other costs include professional development to improve the skills of teachers who prepare students for the exam and the costs associated with the multiple opportunities students have to retake the exam (Center on Education Policy, 2004; Hoff, 2006).

For students with disabilities, as well as other students, the consequence of failing the exit exam means the denial of a high school diploma or remedial classes to improve the student's chance of passing the next retake of the test. If students repeatedly fail the high school exit exam, they may need to remain in school longer to meet requirements for a diploma. Students may have a lowered self-esteem, or they may drop out of high school if they cannot pass the exam (Christenson, Decker, Triezenberg, Ysseldyke, & Reschly, 2007; Johnson & Thurlow, 2007).

Even though there has been an identified association between state exit exams policies and student drop-out, California is one of 22 states that require high school students to complete an exit examination to receive a high school diploma (Center on Education Policy, 2007). After determining state proficiency standards were too low, California decided to set higher standards for high school graduation. The decision resulted in the adoption of a high school exit exam (California Department of Education, 2008). The California high school exit exam (CAHSEE) is a graduation requirement that was first authorized by state law in 1999. The purpose of the CAHSEE is to significantly improve student achievement in public high schools and to ensure that students who

graduate from California's public high schools can demonstrate grade-level competency in reading, writing, and mathematics.

All students are required to take the CAHSEE for the first time in tenth grade. Students who do not pass one or both parts of the exam in grade ten have up to seven opportunities to retake the test by the end of the 12<sup>th</sup> grade. School districts are required to provide additional instruction to assist students who do not pass the exam. Students who have not passed one or both parts of the CAHSEE by the end of grade twelve are entitled to receive intensive instruction and services for up to two consecutive academic years after completion of grade twelve or until they pass both parts of the CAHSEE, whichever comes first. In order to pass the CAHSEE, a scale score of 350 or higher is required on each part. Students who pass one part of the CAHSEE only need to retake the part not passed (California Department of Education, 2008d).

**Gaps in passage and graduation rates.** Along with the high costs of exit exam policies is the large gaps in passage rates. The Center on Education Policy (2007) reported that all states with exit exam policies exhibited gaps in pass rates between various groups; however, the largest gap exists between students with and without disabilities. This gap could range up to 44 percentage points in mathematics and reading and language arts. In California, out of all students in the 10<sup>th</sup> grade who took the exit exam for the first time in 2006, 24% failed the math section and 23% failed the English language arts section. However, 69% of students classified as special education (students with disabilities) in the class of 2006 failed both sections on the test (California Department of Education, 2008a).

In states that report passing rates by student subgroups, African American, Hispanic, low-income, non-English speaking, and students with disabilities on average

scored 30 to 40 percentage points below Asian and White students. For example, in Massachusetts, Asian American and White students were twice as likely as Latino students and one and three-quarter times as likely as African American students, to pass the state math exam on the first try (Center on Education Policy, 2002). Other states show a wide disparity in initial pass rates as well. For example, in 2006, African American students in Indiana scored 38 percentage points lower, Latino students scored 25 percentage points lower, and Native American students scored 15 percentage points lower than White students in mathematics. In addition, African American students scored 34 percentage points lower, Latino students scored 30 percentage points lower, and Native American students scored 13 percentage points lower than white students in reading and English language arts, respectively (Center on Education Policy, 2007).

Pass rates on the CAHSEE for students with disabilities in California is low, approximately 37% in 2009 (California Department of Education, 2009). This is of concern because exemptions for students with disabilities were permitted for those who had completed all other graduation requirements until December 31, 2007 (California Department of Education, 2007). In 2008 all students who passed the exit exam on the first try equated to 65.1%. During that same year, students designated as special education had a pass rate of 20.9% (California Department of Education, 2008a). The California Department of Education reports these numbers as the most accurate ever collected because for the first time it relied on individual student identification numbers.

A majority of students do successfully complete the exam by the time they are ready to graduate; however the data can be misleading. The Civil Rights Project (Planty, Hussar, Snyder, Provasnik, Kena, Dinkes, Kewal-Romani, & Kemp, 2008) suggests that the graduation rates with a diploma reported by the National Center for Education

Statistics (NCES) are flawed data. Schools often report students who never receive diplomas as successfully transferring to other schools or data on dropouts are often unavailable. For example, in 2002, Indiana and Ohio reported approximately 98% of test-takers eventually passed the exit exam and received a diploma, but the numbers of students on whom the rates were based did not include students who dropped out of high school, repeated their senior year, moved away, or were excluded from testing because of disability or language status (Center on Education Policy, 2002). As a result, states' graduation rates may be inflated. For example, in California, the reported increase in graduation rates of 71.6% to 74.6% during the years 2001-2005 did not account for dropouts or students who left the state of California (Planty et al., 2008).

Although there are discrepancies in state reported graduation rates, many American students are not graduating with a high school diploma. In 2004, 70% of American students finished high school with a diploma (Swanson, 2008). In urban areas, the percentage is lower. The average high school graduation rate with a diploma for urban school districts across the country is approximately 60%. The most extreme cases have been in Baltimore, Cleveland, Detroit, and Indianapolis where fewer than 35% of all students graduate with a diploma (Swanson, 2008). In 50 of the largest U.S. cities, approximately 52% of students graduate with a high school diploma (Swanson, 2008). Among certain populations (e.g. racial and ethnic minorities, males, and students with disabilities) the rates are approximately 15% lower.

In 2007, California reported graduation rates with a diploma of approximately 68%. During that same year dropout rates were reported to be approximately 21%. The other 11% exited high school without a diploma. The latter group consists of students who transferred to a private school, left the state, took the General Educational

Development Test (GED) to earn a California High School Equivalency Certificate, or students with disabilities who received a certificate of completion or certificate of attendance in lieu of a diploma (California Department of Education, 2008).

Students with disabilities have a lower graduation rate compared to students without disabilities. The National Center for Education Statistics, in 2008, reported that in 2006, the percent of students with disabilities graduating with a high school diploma nationwide was slightly more than 50% whereas the pass rate for students without disabilities was approximately 70%. The percentage tended to be lower than the national rate when broken down by classification of the disability. The largest classification of students with disabilities is specific learning disabilities (60%); yet 62% graduate with a high school diploma (Planty et al., 2008). The rates are even lower for other exceptionalities: 37% for students with mental retardation, 43% for students with emotional disturbance, and 44% for students with multiple disabilities receive a high school diploma.

As a result of the federal accountability system and low pass rates among first-time test takers, several studies have been conducted on high school exit exams. In a study conducted by Cornell, Krosnick, and Chang (2006) results of surveys from 911 out of approximately 1,500 students who received inaccurate notices of failure on the exit exam indicated that more than three fourths of the students reported adverse emotional reactions to failing the Minnesota exit exam. Students were distressed when told they failed the high stakes test.

A survey of approximately 450,000 students' perceptions of the CAHSEE was conducted by the Human Resources Research Organization for the California Department of Education (Wise et al., 2005) to indicate trends in preparation for the exam and

subsequent plans. Results suggested that students perceived the exam as important, and test questions addressed topics that had been covered in coursework. Economically disadvantaged students or those who did not pass the tests perceived the test as more important to them and made an extra effort besides regular course work to prepare for the tests. They also indicated uncertainty about their high school graduation, nervousness when taking the test, unfamiliarity with the test questions, and most test takers felt the test questions were more difficult than their regular course work. Approximately 10% of examinees reported they had not taken the courses that covered the topics, and 20% reported difficulty with the topics when they were taught. Most test takers indicated that the tests were no more difficult than their course work. Yet, the failure rate on the exams remained high for first time examinees. In the current study failure rate of first time examinees will be explored. This is important because first time examinee scores are counted in California's AYP accountability report to the federal government and the state's own API accountability reporting.

A negative reaction towards exit exams has prompted research on predictive factors related to state exit exams. Research has indicated a significant relationship exists between poor achievement at an early age and later poor achievement, grade point average and achievement on standardized tests, discontinuity and disengagement in school, academic history, and absenteeism (Nichols, 2003). Students have a difficult time being successful on any standardized test if they had not attended school to acquire the knowledge. Lower socioeconomic status, to a lesser extent, is another prohibitor for success on standardized tests.

Recent research has been conducted in California to determine factors that may contribute to the success or failure rate on the CAHSEE. In a study conducted in

Southern California, Zau and Betts (2008) found that the strongest factor for predicting success on the CAHSEE was academic Grade Point Average (GPA) in freshman year, attendance, and classroom behavior. A one-point increase in GPA or a large jump in math scores on the California Standards Test (CST) in 9<sup>th</sup> grade was also associated with a higher probability for passage of the CAHSEE.

A relationship between school level factors and academic factors also suggested that taking Algebra 1 or a higher level of mathematics in 8<sup>th</sup> grade, as well as 8<sup>th</sup> grade CST results, were both strong predictors of 10<sup>th</sup> grade CAHSEE mathematics pass results (Taylor, 2006). Students with disabilities and special needs students, who were also at risk for failing to pass the CAHSEE, were not addressed. As such Taylor (2006) suggested the need to examine relationships in terms of special population variables and CAHSEE results. This study addressed the gap in the literature by investigating differences in student performance on the California High School Exit Exam (CAHSEE) between students with disabilities who were enrolled in a mainstream or a self-contained setting for math.

**Math ability and setting.** Mathematics is important for living in today's society. For students to compete successfully in the global economy, a higher degree of mathematical comprehension is needed. Early access to a rigorous curriculum has been associated with student achievement (Barton, 2004), and algebra has been considered the gatekeeper to academic achievement in high school and in higher education (Jacobson, 2000). To be competitive, the California Department of Education has determined that Algebra 1 is an appropriate grade-level math course for students as early as 8<sup>th</sup> grade. However, many of California's students do not begin Algebra 1 until 9<sup>th</sup> grade or later. Students not enrolled in an Algebra 1 course by 8<sup>th</sup> grade take the General Mathematics

CST, which tests 6<sup>th</sup> and 7<sup>th</sup> grade math standards. When they are enrolled in the Algebra 1 course, they take the Algebra 1 CST. Since algebra is the level of mathematic needed for academic achievement in high school and higher education, a closer examination related to setting for optimal math instruction of students with disabilities is needed.

A strategy used to increase the performance of students with disabilities has been to mainstream them into the general education classrooms. The practice of including students with disabilities in the general education classroom has been around for some time. Classroom placement relative to academic performance of students with disabilities has been a debated topic. Although reactions to this practice has varied, it has been found that students with disabilities achieve better outcomes when they have been included in general education classrooms than in a pullout program in mathematics (Rea, McLaughlin, & Walther-Thomas, 2002). Students with learning disabilities in the general education setting when compared to a special education small group setting had comparable scores on the mathematics subtest of a statewide proficiency test. Rea, McLaughlin, and Walther-Thomas (2002) noted that these results have implications for the long-term outcomes of high school graduation and subsequent employment. In contrast, other researchers have found little correlation between setting and academic achievement (Fore III, Hagan-Burke, Burke, Boon, & Smith, 2008). As such, this study investigated whether the setting (general education or special education) had an effect on the pass or fail status on the math section of the CAHSEE.

### **Theoretical Rationale**

The stereotype threat theory (Steele & Aronson, 1995) provided the theoretical rationale for this study. Stereotype threat is the belief or threat of confirming a negative stereotype about a group to which one belongs (Steele, 1997; Steele & Aronson, 1995).

Stereotype threat can affect anyone with a group identity to which a negative stereotype exists. For example, a negative stereotype would be that a student is unintelligent because he or she has a learning disability. The belief is that because a student has a disability, he or she will be judged according to the negative stereotype associated with students with disabilities. Stereotype threat, which can be signaled by the mere recognition or association with a negatively stereotyped group, may be a factor in a student's low academic performance (Steele, 1997). Due to the high personal stakes connected to the results of the CAHSEE, the threat of confirming a negative stereotype may factor into the low first time passage rate of students with disabilities.

Stereotype threat describes the possibility of confirming a negative stereotype, not the actual experience. Thus, the stereotype does not have to be real. However, if the threat is experienced in the midst of a classroom presentation or test, for example, the emotional reaction it causes may interfere with performance. This threat can hinder a person's intellectual performance when the negative stereotype relates to the person's intellectual or academic ability. When an individual from a stereotyped group performs a scholastic or intellectual task, the threat of confirming or being judged by a negative societal stereotype about the group's intellectual ability and competence interferes with the individual's intellectual functioning on standardized tests. In other words, the individual may experience performance related stress or anxiety which can negatively affect the individual's performance when the exam is most important to the individual (Osborne, 2001; Steele, 1997), as is the case with the high school exit exams.

Steele and Aronson (1995) first used stereotype threat to explain why black college freshmen and sophomores performed worse than their white counterparts on a standardized test when their race was emphasized. The negative stereotype was that black

students were intellectually inferior to white students because of their race. The results indicated that African American students' standardized test performance, relative to their white counterparts, was depressed because of the African American students' vulnerability to judgment about their group's intellectual ability based on negative stereotypes. Since then, other researchers have attributed gaps in achievement levels to stereotype threat. In particular, a recent study by Reardon, Atteberry, Arshan, and Kurlaender (2009) examined the effects of the exit exam requirement on the outcome of students with low prior achievement levels. Their findings suggested students with low 10<sup>th</sup> grade academic skills, minority students, and females were most affected by exit exams because of a stereotype-threat induced fear of failing the test or concern about proving a negative stereotype.

Steele and Aronson's stereotype theory provides a linkage between academic history and standardized test performance. Several studies have suggested that underachievement on academic tasks, poor test performance in academic environments, and lower test performance on statewide standardized tests are all consequences of stereotype threat for the individuals identified within the stereotyped group (e.g. Cole, Matheson, & Anisman, 2007; Good, Aronson, & Harder, 2008; Good, Aronson, & Inzlicht, 2003; Keller, 2007; Neuville & Croizet, 2007). For instance, 138 junior high school students' math and reading performance on a statewide standardized achievement test, used in promotion or retention of students, revealed that the females' (45%) performance was lower under stereotype threat compared to the males' performance. The gap in performance was reduced with the reduction of the threat (Good, Aronson, & Inzlicht, 2003). Stereotype threat has also been suggested to describe weak student

performance on tests that were viewed as free of cultural bias and not reliant on reading skills or background knowledge (Brown & Day, 2006).

Stereotype threat has transitioned beyond the traditional group identity factors (e.g. race and gender). For example, an older adult's faltering memory will confirm stereotypes about the elderly (Steele, 1997). This study expanded on group identity to include student with disabilities. The present study augments Steele and Aronson's discussion of stereotype threat of intellectual ability by including special education identity status as a stereotyped category.

Stereotype threat guides this study because it has been seen as a factor in explaining achievement gaps in standardized testing and applies to any group membership where underperformance could confirm a stereotypical expectation. Students are usually identified or recognized as a student with a disability when taking the CAHSEE. Given that students' with disabilities prior academic performance has been lower than students without disabilities and the threat of being negatively stereotyped because of the student's group identity, especially if the student does not pass the first time, may be in relation to the students' CAHSEE performance. The more a test outcome is associated with a clear statement about success or failure, not unlike the exit exam, the more it may impact a student's performance (Brown & Joseph, 1999). Thus, the study explored whether 10<sup>th</sup> grade students with disabilities placed in mainstream or self-contained settings for mathematics instruction differ in relation to a stereotype threat, anxiety, and their pass/fail status on the math section of the CAHSEE. In addition the study will examine whether the relational patterns among math achievement (e.g. course grade, CST mathematic scores), anxiety, stereotype threat, and CAHSEE mathematic scores are similar for both groups.

## Research Questions

The primary purpose of this study was to examine the differences in student performance on the California High School Exit Exam (CAHSEE) between 10<sup>th</sup> grade students with disabilities who were enrolled in a mainstream or self-contained setting for math. Students with disabilities have the lowest percentage rate of passage compared to whites and other minority groups. As such, the exam has a tremendous impact on the post high school achievement of students with disabilities because the high school exit exam determines whether a student will receive a high school diploma.

The following research question(s) were explored in this study:

1. When controlling for prior math ability (as measured by Algebra 1 grade, 8<sup>th</sup> and 9<sup>th</sup> grade CST scores on Math), is there a significant mean difference in the pass/fail status on the mathematics section of the CAHSEE for 10<sup>th</sup> grade students with disabilities enrolled in mainstream or self-contained settings for math instruction?
2. When controlling for prior math ability (as measured by Algebra 1 grade, 8<sup>th</sup> and 9<sup>th</sup> grade CST scores on Math), is there a significant mean difference in stereotype threat between 10<sup>th</sup> grade students with disabilities enrolled in mainstream or self-contained settings for math instruction?
3. When controlling for prior math ability (as measured by Algebra 1 grade, 8<sup>th</sup> and 9<sup>th</sup> grade CST scores on Math), is there a significant mean difference in anxiety between 10<sup>th</sup> grade students with disabilities enrolled in mainstream or self-contained settings for math instruction?

## **Definition of Key Terms**

*Academic Performance Index (API)* - used to measure the academic performance and growth of schools. The API is one component of California's definition of Adequate Yearly Progress (AYP), required under the No Child Left Behind (NCLB) Act of 2001. Only a student's first time attempt of the CAHSEE in the spring of their 10<sup>th</sup> grade school year counts towards the API.

*Adequate Yearly Progress (AYP)*- is a measurement defined by the United States federal No Child Left Behind Act that allows the U.S. Department of Education to determine how every public school and school district in the country is performing academically according to results on standardized tests.

*Anxiety* - the response to stressful situations based on the extent to which each individual perceives the specific situation as psychologically dangerous or threatening, and is greatly influenced by each individual's past experience.

*California High School Exit Exam (CAHSEE)* - a high school graduation requirement or a diploma in the state of California. Students must pass the exam, and met district requirements to receive their high school diploma.

*California High School Graduate*- a student who left school after meeting all state and local high school graduation requirements, including passing the California High School Exit Exam (CAHSEE), and received a standard high school diploma.

*California Standards Test (CST)* - a state standard test covering English-language arts, mathematics, science, and history-social science design specifically to assess student performance on California's content standards. The standards are specific to grade and course. Grades 2-11 receive content standards test in both language arts and mathematics; grades 8, 10, and 11 receive content standards test in history-social science; and grades 5,

9, and 11 receive content standards test in science. The mathematics standards test is course specific. The “General Mathematics” standards test is specifically for seventh graders, but can be taken by students with disabilities not enrolled in an Algebra I course level or higher.

*Certificate of Achievement-* given when a student has completed all coursework requirements for graduation, but has failed to receive a passing score on one or both parts of the exit exam.

*Certificate of Attendance-* given when a student has not passed the exit exam and has not completed all coursework requirements for graduation, but has attended high school for at least four years.

*High Stakes Test-* the mechanism by which states measure whether students are meeting state standards.

*No Child Left Behind-* an act of 2001, designed to improve student achievement. The passage of No Child Left Behind reauthorized the Elementary and Secondary Education Act (ESEA), the principal federal law affecting education from kindergarten through high school.

*Setting-* In this study setting is defined as a Mainstream Class or Self-Contained Class:

*Mainstream Class-* a classroom on the school campus where the delivery of instruction is conducted by a general education teacher. The instruction is provided in a large group setting where the majority of students are considered students without disabilities.

*Self-Contained Class-* a separate classroom on the school campus where the delivery of instruction is conducted by a special education teacher. The instruction is

provided individually or in a small group setting where all students in the class are considered students with disabilities.

*Stereotype Threat*- refers to being at risk of confirming a negative stereotype or expectation about one's group. If the threat is strong enough, it can impair an individual's intellectual and academic performance.

*Students with Disabilities*-a student who needs special education and related services as a result of mental retardation, hearing impairments (including deafness), speech or language impairments, visual impairments (including blindness), serious emotional disturbance, orthopedic impairments, autism, traumatic brain injury, other health impairments, or specific learning disabilities (Individuals with Disabilities Education Act (IDEA), Public Law, 101-476).

## **CHAPTER TWO**

### **LITERATURE REVIEW**

The high school exit exam is a graduation requirement that must be successfully completed before a student receives a high school diploma. Graduation from high school with a diploma is an essential component for accessing further training, education, military, or career options (Dunn, Chambers, & Rabren, 2004). All students are required to take the exit exam, including students with disabilities and English language learners, beginning their sophomore year of high school. First time failure rates are high among certain subgroups; however, students with disabilities fail at a much higher rate than any other subcategory of students, placing them in danger of not receiving a high school diploma. Unfortunately, little has been done to empirically investigate the problem of high first time failure rates of students with disabilities on the California High School Exit Exam mathematics section.

Presented in this chapter is a review of relevant literature on major factors identified as contributors to the failure of secondary students with disabilities on the first administration of the California High School Exit Exam mathematics section during their sophomore year. The chapter is organized in three sections: mainstreaming and self-contained settings, high-stakes testing for students with disabilities, and stereotype threat. The chapter concludes with a summary of the literature reviewed.

#### **Mainstreaming and Self-Contained Settings**

As more students with disabilities are included in general education classrooms, the relation between placement and outcomes, or which placement is optimal for educating students with disabilities, has been a critical issue. The concern is that students

with disabilities are not achieving at optimal levels in more inclusive settings because they require specialized instruction. Likewise, many students with disabilities in specialized settings fall short of their academic potential because they lack access to the general curriculum. In response to this issue, several researchers have investigated differences in setting for students with disabilities and their academic performance.

Rea, McLaughlin, and Walther-Thomas (2002) conducted a study that compared the performance of middle school students with learning disabilities (LD) who were served in inclusive classrooms (mainstream) with similar students served in pullout special education programs (self-contained). Students were compared across academic achievement, daily school attendance, and disciplinary infractions.

The researchers utilized archival qualitative and quantitative data from all 8<sup>th</sup> grade students with LD in two middle schools within the same suburban school district in the southeast. The two middle schools were matched on as many factors as possible so that the ability to attribute any differences in outcomes observed to setting would be increased. The students for the study were identified by a computer search of the December 1 Federal Child Count conducted by every school. To verify previous enrollment in special education, the prior year's child count records were used. Two groups of students were analyzed. Thirty-six students were from school one and received special education services through an inclusive support model (mainstream) and twenty-two students were from school two where they received special education services through "pullout" model (self-contained classrooms).

Similar to this study, the data were collected to analyze the relationship between placement of students with LD and school performance. Student and family data were clustered as demographic variables while the students' grades, behavior, and attendance

was clustered as outcome variables. Program information was collected about the school district and each of the two middle schools (e.g., program, IEP, and teacher variables). Once identified, data were retrieved from the December 1 Child Count records, Individualized Education Plan (IEP), special education eligibility records, individual student evaluation reports, class schedules, attendance records, discipline records, report cards, and student scholastic records. The information regarding chronological age, gender, race, socioeconomic status, education level of the mother, disability category, estimated IQ at the most recent triennial evaluation, years receiving special education services, years enrolled in the present school district, report card grades, standardized test scores, state proficiency test scores, disciplinary actions, and daily school attendance were compiled for each student at the end of the 8<sup>th</sup>-grade year. The service delivery models for students with LD were verified through teacher planning documents, supervisor observation notes, students' IEPs, teacher and student schedules, and team meeting minutes.

School one served students in grades 6 through 8 with a teaming model. During the period of investigation, 1994-1996, three four-person teams of general educators served the students. The model in school one for implementing inclusive special education services was based on team teaching and collaborative planning. General and special education teachers co-taught four periods per day and had one period of individual planning and one period of team planning. School two also served students in grades 6 through 8 in a teaming model, but students with disabilities in school two received their services from two teachers each year in a separate setting from their general education teachers. Students in school two received no in-class special education services. The four core courses (language arts, mathematics, science, and social studies)

were taught in general education classrooms by general education teachers working alone. Pullout services were scheduled during elective periods. Students forfeited either one or both of their elective classes to receive special education services. During pullout sessions one special education teacher worked with a small group of identified students to remediate academic weaknesses or to assist with completion of assignments for general education core classes.

Three indicators of student outcomes were measured. First, academic achievement was based on final course grades in the 8<sup>th</sup> grade language arts, mathematics, science, and social studies, curricula. Second, the student's standard scores on the reading, mathematics, science, and social studies subtest of the Iowa Test of Basic Skills (ITBS). Third, the students' highest scores, pass/fail rates, number of administrations, and nonstandard administrations on the reading, mathematics, and written language subtest of the state's academic proficiency test the Literacy Passport Tests (LPT) were measured. Course grade was a teacher determined measure of student achievement in each course. Course grades were reported in letters based on a district-approved point system (A = 94-100, B = 85-93, C = 75-84, D = 69-74, F = 0-68). The final course grades in language arts, mathematics, science, and social studies were collected from student report cards and converted to a standard four-point scale.

The state proficiency tests are designed to measure mastery of state learning objectives in reading, language arts, and mathematics. Students had to pass all three subtests to obtain the Literacy Passport, a requirement for graduation from high school with standard or advanced diplomas. Students who did not pass all portions of the tests were provided with repeat opportunities to take the failed portions. This is similar to the procedures for the California High School Exit Exam (CAHSEE). School behavior was

defined as actions (e.g. disruption, disobedience, fighting) that resulted in an in-school or out-of-school suspension, and school attendance was determined by the number of days per school year each student was absent from school.

To establish comparability of the groups in terms of their chronological age, gender, ethnicity, socioeconomic status, mother's education level, estimated cognitive abilities, years receiving special education services, and years in the current school district *t*-tests or chi-square analyses were conducted. Results indicated that the majority of students in both settings were Caucasian with no significant difference in ethnicity. The groups were similar in terms of students receiving free or reduced-fee lunch, ethnicity, and mother's educational level with most having at least a high school diploma and approximately half attended college. No significant differences were found for socioeconomic status or cognitive abilities. Students were comparable on the mean number of years that they received special education services (approximate 6 years) and enrollment in the school district (approximately 5 years).

Program variables, including number and nature of IEP goals and objectives, degree of classroom accommodation, and amount of special education service delivery for students in the two groups were targeted. The IEP goals and objectives data relating to number of accommodations and amount of special education services students received were collected from a review of students' IEPs and cross-referenced with the students' class schedules. A panel of doctoral students were trained in review procedures and provided copies of the school district's curriculum and coding forms. Each graduate student reviewed approximately one third of each school's IEPs. Interrater reliability was established at .92 by distributing 10% duplicates. Coders classified goals and objectives according to standard curriculum used by the district, remedial basic skills, thematic

units, learning strategies, study skills, affective/behavioral skills, or vocational/career skills.

Results related to accommodations and amount of special education services suggested programs differed primarily in type and intensity of special education services delivered, skills addressed, implementation of classroom accommodations, teacher consultation, and instructional models. IEPs developed at school one contained significantly more instructional goals than those at school two. Significant differences were also found in two categories of goals, those focused on general education curriculum and remedial skills. School one's IEP committees established goals for students that reflected school district learning expectations for all 8<sup>th</sup> graders. At school two, the committees focused on academic deficits and established goals to remediate those areas. IEP goals that addressed learning strategies and student behavior were not significant. Specific performance expectations revealed that IEPs written for inclusive services contained significantly more objectives than those written for pullout service delivery.

Regarding classroom accommodations, analyses revealed statistically significant differences in patterns of implementation. More accommodations for students in school one were discovered compared to school two. Moreover, the accommodations were intended for use in general education classrooms and fell into three categories: instruction, assessment, and behavior. There was also a significant difference in implementation of accommodations to address student behavior. Although students with LD served in inclusive classrooms earned significantly higher grades in all four areas of academic instruction, statistical analysis of data on performance on the reading,

mathematics, and writing subtest of the state proficiency test revealed no significant differences between students with LD receiving inclusive and pullout services.

Analysis of the ITBS results was mixed. Fifty-four students participated in ITBS testing in their 8<sup>th</sup> grade year, and four students were exempted. Results indicated students with LD who received inclusive special education services achieved higher standard scores on the language and mathematics subtests than students with LD who received pullout special education services. The two groups earned similar mean scores on the reading comprehension, science, and social studies subtests. Regarding behavior, students with LD in inclusive classrooms did not experience more in-school or out-of-school suspensions than students in pullout programs, but attended more days of school than those in pullout programs.

Similar to Rea et al. (2002), the present study investigated students with disabilities who were enrolled in mainstream settings or self-contained settings for mathematics instruction and their performance on the CAHSEE. Rea et al. (2002) concluded that self-contained classrooms for academic instruction does not equate to improved scores or pass rates on standardized tests. While the field of special education evolved to serve students with specific needs, data on self-contained settings for students with LD revealed non-satisfactory in terms of school achievement or long-term benefits (Rea, McLaughlin, & Walther-Thomas, 2002). The study suggests that students with disabilities can achieve academic success in general education classrooms. This implies that students in a mainstream setting for math instruction will have a better chance at passing the CAHSEE than students taught in a self-contained setting. A weakness of the study was the small sample size and the fact that the study was conducted in one small school district. In addition, Rea et al. (2002) used a sample from middle schools which do

not have the same impact of high-stakes tests as would a sample from a higher grade level. In addition, the study limited the sample to one category of disability. The current study used more than one school district to increase the sample size which may assist in the generalizability of the study. The categories of disability will be expanded to include all students with disabilities who are subject to the exit exam requirement. Further, the current study examined high school test performance in a significantly higher stakes environment where a high school diploma, an essential social marker of achievement and passage, is at stake. Additionally, the present study focused on achievement in mathematics. Thus, the present study may have broader generalizability for the field of special education.

Other researchers have also examined classroom placement, inclusive versus non-inclusive or mainstreaming versus self-contained, relative to the academic performance of students with specific learning disabilities but with high school students. Fore III, Hagan-Burke, Burke, Boon, and Smith (2008) conducted a descriptive exploratory study to examine differences in academic achievement based on classroom placement for students with specific learning disabilities (SLD) in secondary classroom settings. Fifty-seven high school students with SLD, forty-two males and fifteen females, from two suburban high schools in the southeastern United States were participants in the study. Out of the participants, 19 were in 9th grade, 18 in 10th grade, 13 in 11th grade, and 7 in 12th grade. The ethnic backgrounds of the students included 50 Caucasian, 5 African-American, and 2 Hispanic-American students. Eighty percent of the participants were reported to have reading disabilities and 20% had math disabilities. There were no instances of comorbidity (i.e., diagnosis of both reading and math disabilities). Approximately 10% of

the student participants had an additional diagnosis of attention deficit hyperactivity disorder (ADHD).

The participants received special education services in inclusive and non-inclusive settings, varied by the number of inclusive classes that each was scheduled. Inclusive classes were those that were offered to general education students and taught by a general education teacher. Most inclusive classes contained approximately 25 students with no more than 20% identified as having a disability. Some (e.g., mathematics and literature classes), but not all of the inclusive classes also had a special education teacher. Conversely, classes deemed as non-inclusive were those that were taught by a special education teacher that occurred in settings other than a general education classroom. Non-inclusive classrooms contained no general education students and were only taught by special education teachers. The students were asked to complete the Grade Level Test Short Form of the *Multilevel Academic Survey Test* (Howell, Zucker, & Morehead, 1985) because of its high level of agreement with the states' standardized test for reading and mathematics. The *Multilevel Academic Survey Test* (MAST) consisted of two 20-item multiple choice maze tasks and 24 math computation items. The MAST was intended for use by school personnel to make decisions about student performance in mathematics and reading. The test can be used by teachers in general classrooms with normally achieving students, but was primarily intended for those educators who instruct and assess students exhibiting academic difficulties.

Trained graduate students administered all measures to student participants. The graduate students established a testing schedule with the participating school and administered the MAST to groups of students in accordance with the test's guidelines. The protocol provided by the MAST administration manual was followed for both the

reading and math subtests. On the reading portion of the MAST, the students were given five minutes to complete the first maze task. Then the students were instructed to stop and turn to the next maze task. Again, students were given five minutes to complete the second maze task. Next, students were instructed to complete the arithmetic problems on their answer sheet. After twelve minutes, they were asked to stop and the administrator collected the tests.

Descriptive statistics and significance differences were reported for gender and grade level, number of general education classes attended, and type of placement. No significant differences were observed for reading or math across grade levels as well as MAST reading or math across class schedules. With regards to type of class taken, significant differences on the MAST reading test were observed between those students who took the general education literature class for reading and those who did not with students in the general education class performing better than students in the self-contained class. However, the effect size for those differences was small. Overall, with the exception of one comparison, no statistically significant differences in the academic performance of students with SLD for reading or math were found. Fore III, Hagan-Burke, Burke, Boon, and Smith (2008) reported findings suggest little support for one class placement over another.

As with Rea, McLaughlin, and Walther-Thomas (2002), Fore III, Hagan-Burke, Burke, Boon, and Smith's (2008) small sample size impeded the ability to make generalizations to a larger population. The researchers conclude that students with disabilities may achieve similarly on standardized tests whether placed in a mainstream or self-contained setting for academic instruction. This implies that any differences observed on standardized test scores may be the result of something other than

instructional setting. The present study investigated this implication by using a larger sample size and controlling for academic ability.

Delivery of education, in the best possible environment for learning, to students with disabilities has been and continues to be a goal of special education (Kavale & Forness, 2000). Delivery typically occurs in a mainstream classroom setting or self-contained setting. Advantages of self-contained classrooms include smaller class size, curriculum focused on teaching life skills, and more opportunities for individualized instruction to meet the needs of students with disabilities (Signor-Buhl, LeBlanc, & McDougal, 2006). In contrast inclusive classrooms provide for the needs of students with disabilities within the general education environment. Which placement is best with regards to the academic growth and social development of students with disabilities is unclear. Signor-Buhl, LeBlanc, and McDougal (2006) conducted a quasi-experiment to evaluate the academic outcomes of children in self-contained versus inclusive models of special education programs. The researchers examined the comparison of the academic progress of students served in self-contained and inclusion programs and document the results generated.

The study was conducted in a midsized urban school district in Upstate New York. District records were reviewed for students in fourth-grade inclusion classrooms compared to a group of students from self-contained classrooms within the same district. Participants were excluded if they had not attended their self-contained or inclusive education program at least two years prior to the study. Participants with significant disciplinary problems were also excluded from the study to avoid possible confounding variables related to student conduct. Males represented 47 % of the self-contained group

and 64% of the inclusion group. The ethnicity of the students in both groups was White (35%), African American (51%), and Other (14%).

Data were collected from a review of class lists, cumulative folders, and databases that contained students' scores on district- and state-wide assessments given by the school district between 1997 and 2004. Intelligence test scores were used to control for cognitive differences between the inclusive setting and the self-contained setting groups. Achievement measures were used to determine academic outcomes for both the inclusion and self-contained groups. Scores from several intelligence tests were used: The *Wechsler Intelligence Scale for Children, Third Edition* (WISC-III; Wechsler, 1991a), the *Stanford-Binet Intelligence Scale, Fourth Edition* (SB-IV; Thorndike, Hagen, & Sattler, 1986), the *Kaufman Brief Intelligence Test* (KBIT; Kaufman & Kaufman, 1990), the *Leiter International Performance Scale-Revised* (Leiter-R; Roid & Miller, 1997), and the *Slosson Intelligence Test-Revised* (SIT-R; Slosson, 1996).

The scores from the state mandated high-stakes assessment of English and Language Arts (ELA) skills for all fourth-grade students was used as a measure of achievement. When available, data from individually administered achievement tests were collected: the *Wechsler Individual Achievement Test* (WAIT; Wechsler, 1991b), *Wechsler Individual Achievement Test, Second Edition* (WAIT-II; Wechsler, 2001), *Woodcock-Johnson Revised Tests of Achievement* (WJ-R; Woodcock & Johnson, 1989), *Woodcock-Johnson III Tests of Achievement* (WJ-III; Woodcock, McGrew, & Mather, 2001), *Kaufman Test of Educational Achievement* (K-TEA; Kaufman & Kaufman, 1998), *Wide Range Achievement Test-3* (WRAT-3; Wilkinson, 1993), *Test of Early Mathematics Ability, Second Edition* (TEMA-2; Ginsburg & Baroody, 1990), and the

Test of Early Reading Ability, Second Edition (TERA-2; Reid, Hresko, & Hammill, 1989).

Descriptive statistics were used to investigate the demographics of the two groups with regard to age, gender, ethnicity, Full Scale IQ, free or reduced lunch status, and average time spent in inclusive or self-contained settings. An analysis of covariance (ANCOVA) was used to determine significant differences in achievement between the two groups in reading and math with the students' full scale IQ as the covariate. Effect sizes were used to examine the amount of an effect the programs had on outcome measures. To determine the effect sizes, average  $z$  scores were calculated for each group and then the control score was subtracted from the treatment score which provided a change in the  $z$  score.

The average age of the participants showed a slightly higher mean for the self-contained group (11.42 years) than the inclusive setting group (10.84 years). Students in the self-contained group had a mean Full Scale IQ score of 76.19 whereas students in the inclusion group displayed a mean score of 79.32. Percentages of students in the self-contained (89.5%) were slightly higher than the inclusion group (87.1%). The time students spent in their placement was similar for both groups.

After controlling for IQ, results indicated that students in inclusive classrooms performed significantly better on individual measures of reading achievement producing a moderate effect. Students in the inclusive classrooms performed significantly better on the ELA assessment than students in self-contained classrooms, but similar to students who were in self-contained classes on individual measures of math achievement. A small positive effect for math achievement was found for children in inclusive settings.

Students in inclusive settings performed slightly better on individual measures of math achievement, but not to any significant degree.

The results indicated students with disabilities who were educated in inclusive settings will achieve at the same rate or somewhat better than students with disabilities educated in self-contained setting. A limitation of this study was the fact that the researchers used archival school data. Although the findings suggest support of inclusive classroom when educating students with disabilities, the extent to which setting is relevant to high-stakes testing achievement in high school with students with disabilities was not presented in this study. In the current study, the researcher compared the settings and used an analysis of covariance to determine whether there are significant mean differences in achievement between students with disabilities who receive math instruction in mainstream or self-contained settings. Instead of controlling for IQ the study used prior math ability as the covariate.

The current section reviewed literature relative to mainstream and self-contained instructional settings for students with disabilities. A review of research comparing the academic performance of students with disabilities at both the middle and the high school levels within school and on standardized tests was done. The research wanted to determine which setting was most beneficial for the students' achievement and performance. The results were mixed.

Understanding the relationship of setting and performance for high school students with disabilities is critical to improving their performance on high-stakes standardized tests. Since high school students with disabilities must pass the exit exam before they can receive a diploma, it is important that special education educators and stakeholders understand this relationship. However, current empirical research has not

offered findings relevant to the role setting in the success or failure of students with disabilities on high stakes high school exit exams. What has been identified however is that students with disabilities in inclusive environments for academic instruction have similar or slightly higher performance on standardized tests than students in self-contained environments.

### **High Stakes Testing for Students with Disabilities**

Several issues have surfaced regarding the participation of students with disabilities in state- and district-wide assessments. The first issue was the lack of comprehensive data regarding the participation of students with disabilities in large-scale assessments and the corresponding impact on high school completion rates. The second issue was the limited amount of empirical evidence concerning the effects of high-stakes tests for individual students, especially those with disabilities (Ysseldyke et al. 2004). A third issue related to indicators for success or failure on high school exit exams.

Given that students' academic experiences are heavily driven by the NCLB accountability measures and the need to meet adequate yearly progress proficiency rates, Zhang, Katsiyannis, and Kortering (2007) conducted a study to examine the performance of students with disabilities on End-of-Course (EOC) exams. Zhang's et al. (2007) investigation compared student performance in several ways: comparing pass rates of students with and without disabilities to test differences between the two groups, identifying the yearly progress made by students with disabilities in order to draw conclusions regarding the worthiness of including students with disabilities in statewide exit exams, and linking student performance to local education agency (LEA) socioeconomic status (SES), as measured by the percentage of students receiving free or

reduced price lunch, to examine the relationship between student pass rates and poverty levels.

In an effort to improve student outcomes in the 1990s, North Carolina implemented a statewide (EOC) exam system. In the system, all students had to pass state exit exams in five subjects, including Algebra 1, Biology, Economics and Civics, English 1, and Physical Science give within the last 10 days of a given semester. Data for the study were obtained from the North Carolina Disaggregated Performance Data published by the Public Schools of North Carolina in 2005. The data collection covered a 4-year period from 2000-2004 in which student pass rates for each year in Algebra 1, biology, economics and civics, English 1, and physical science, listed by LEA and aggregated across the entire state were obtained. The pass rates were also listed according to disability status: students without disabilities, students with disabilities as a whole, and students with specific categories of disabilities. For the purpose of the investigation, data were extracted from four subjects: Algebra 1, Biology, English 1, and Physical Science to compare performances between students with and without disabilities.

In North Carolina, students must pass each of the subjects and the respective EOC exam to earn an initial level standard diploma called Career Prep. Two other standard diploma options, College/Tech. Prep and College/University Prep, were also available and required students to pass additional subjects. Furthermore, students had to pass the subjects in the 9th or 10th grade, or it became highly unlikely that they would have graduated with their age peers.

To compare performance differences, data were extracted for both students with and without disabilities and separately for each of the following groups: learning disabilities (LD), emotional or behavioral disorders (E/BD), educable mental disabilities

(EMD), and other health impairments (OHI). To examine the relationship between poverty and student performance, data about students who received free or reduced-price lunch in each LEA were drawn from a report by the North Carolina Department of Public Instruction in 2003. Lunch data were also used to classify LEAs into SES groups—*high*, *middle*, or *low*—according to each LEA's lunch needy rate or percentage of students on an average day receiving free or reduced-price lunches. Local education agencies with a needy rate below 36% were classified in the *high* group, those between 36% and 65% were classified in the *middle* group, and those with a needy rate above 65% were classified in the *low* group.

The percentages of students passing each of the four subjects in each of the four years were entered into an SPSS file for data analyses. The percentages represented passing rates and were arranged by subject, year, and disability status for all LEAs. The SPSS database contained 96 variables on student pass rates in each of the four subjects, for each of the four years, for each of six disability status groups, and for 120 LEAs. In addition, the database contained the lunch needy rate for each LEA.

Three types of data analysis were conducted. First, statewide data on each subject were graphed by year and disability categories to examine trends in the passing rate in each subject over and across disability categories. Second, descriptive statistics and correlation coefficients were calculated for each of the four subjects for each disability group in 2003-2004 to examine in which subject a particular group performed better (or worse) and to investigate the relationship between pass rates and family SES, measured by the lunch needy rate. Third, repeated measures multivariate analyses of variance (MANOVA) were conducted on each of the four subjects to investigate performance differences among disability status groups (disabilities versus no disabilities) and across

the four years, with comparisons of differences across SES groups. Four separate analyses were performed: each examined one subject (Algebra 1, Biology, English 1, or Physical Science) in relation to two within-subject factors (year and disability status) and one between-subject variable (SES group). Disability status was reduced to two groups, students with disabilities as a whole versus students without disabilities, due to a lack of data on some subjects in some years.

The researchers concluded that students with disabilities who passed the EOC tests range from one-third to one-half that of their peers without disabilities. In addition, students from low-SES schools were less likely to pass the EOC tests and had rates comparable to that of students with disabilities. Results on the Algebra 1 pass rates across student groups revealed four trends. The Algebra 1 pass rate for students without disabilities was significantly higher than for students with disabilities in all four years. Next, students with disabilities as a whole and students with LD showed small but steady increases in Algebra 1 pass rates. In addition, students with EMD demonstrated two increases in the 2002–2003 and 2003–2004 school years. Finally, students with LD and students with OHI had similar pass rates, and both performed better than students with E/BD; in addition students with E/BD had higher passing rates than students with EMD.

Over the four school years, students without disabilities significantly outperformed students with disabilities in all four subjects. Students without disabilities did best in English 1, while students with disabilities did best in Algebra 1. Among the four disability groups (E/BD, LD, EMD, and OHI), students with LD and OHI seemed to perform at the same level, and both groups outperformed students with E/BD and EMD; students with EMD generally had the lowest passing rates in all subjects. Interestingly, students with EMD had relatively high pass rates in Algebra 1 compared to the other

three subjects. All groups showed some improvements in English 1 and Physical Science over the four years. Clear and strong relationships existed between district SES and student performance; that is, students from richer districts outperformed students from poorer districts. When comparing the performances among students with disabilities in different groups, students with LD had the highest pass rates in many instances, although students classified as OHI had similar performances to students with LD.

Zhang et al. (2007) compared students with disabilities and students without disabilities performance on a statewide standardized test in mathematics in which the students must pass to receive their diploma. Since students with disabilities usually are the lowest performing group on standardized test, this study focused on students' with disabilities group performance related to pass rates and placement of math instruction. A major limitation in Zhang et al. (2007) was the use of a single state's data set. Varied degrees of performance existed between districts within a state. The high and low performing districts' students were compared. Therefore, differences in local factors may have had an impact on the results observed in the study. The current study used several similar high schools from similar districts within the state of California to better understand any observed differences found. Similar to Zhang et al. (2007) other researchers have studied the performance of students with disabilities in the high stakes testing environment.

Christenson, Decker, Triezenberg, Ysseldyke, and Reschly (2007) conducted a study to examine the effects of high-stakes assessment for students with and without disabilities. More specifically, the study determined whether differences exist in the consequences of high-stakes assessment for students with and without disabilities. States with high school exit exams were sampled because students in these states would have

encountered at least one high-stakes test in their K-12 education. Eight hundred fifty-six school principals were asked to select from their schools one special education teacher, one general education teacher who taught at a grade level in which the required testing occurred, and one school psychologist assigned to their buildings to complete the survey. Principals received a letter explaining the purpose of the study and survey, guidelines for selecting respondents, and a description of the incentives for respondents who participated in the survey. The principals placed the study materials in the mailboxes of the selected school personnel. Principals who chose to participate in the study returned the survey directly to the investigators in an envelope provided. Teachers who chose not to participate returned the survey to their principals, who then selected another potential participant. Those who responded to the survey were entered into a lottery and given the opportunity to win \$500 in educational materials of their choice.

Eight hundred schools were originally identified from 20 states that required high school exit exams. Twenty elementary schools were randomly selected from each state ( $n = 400$ ) using information from the Common Core of Data, compiled by the National Center for Education Statistics. Twenty middle schools ( $n = 400$ ) were then matched to the school districts of the selected elementary schools. In addition, 56 urban schools were added based on the visual analysis of the school districts represented.

The return rate for this study, 11.6%, was calculated on the basis of the number of schools that returned surveys (i.e., 99 of 856). The study included 249 school professionals, 13% men and 87% women, from 99 schools across 19 states. Of the respondents, 40% were general education teachers ( $n = 97$ ), 37% were special education teachers ( $n = 89$ ), and 23% were school psychologists ( $n = 56$ ). Approximately half of the respondents worked at middle schools or junior high schools, and half worked at

elementary schools. A large portion of the respondents were from rural schools (53%). Others were from urban schools (23%) and suburban schools (24%).

Teachers and school psychologists completed the researcher developed instrument called the *Perspectives of Testing and Grade Promotion Survey*, which examined their perspectives on high-stakes assessment and grade advancement decisions. The survey was developed using two main sources of information: a comprehensive literature review and consultation with a team of school personnel. Key terms grouped on the basis of similarity resulted in nine related themes, all of which focused on increasing student achievement. The themes included holding high expectations and standards for all students, determining how to measure student progress, sharing responsibility for educational outcomes across stakeholders, understanding the outcomes of high-stakes assessment for students and schools, creating clearly specified guidelines for making promotion and retention decisions, preventing early school failure, identifying students with learning difficulties earlier, improving the instructional environment, and increasing the availability of resources. A consultation team, consisting of principals and school psychologists from suburban, urban, and rural school districts, were formed. Information from the themes along with input from the consultation team was used to construct items for the survey. The survey was piloted and minor revisions were made in wording.

The researcher examined whether differences existed in respondents' perspectives on the overall effect of high-stakes assessment for students with and without disabilities. The researchers wanted to identify any specific observable events that were influenced by high-stakes testing for students with and without disabilities. In addition, any differences between students with and without disabilities with regard to the influence of high-stakes tests. The researchers also investigated whether there were differences in how grade

advancement decisions were made for students with and without disabilities and if so, what were the differences that existed. Averages, rank-order or Spearman correlations, and multivariate analysis of variance were used to analyze the data.

The researchers concluded the majority of teachers and school psychologists believed that high-stakes assessment had a positive effect on what students were taught, how students were taught. Also, high-stakes assessment was believed to have a positive effect on the development of alternative instructional strategies. Fewer respondents reported negative effects as a result of high-stakes assessment. Respondents reported instructional changes related to improving student performance had increased along with the monitoring of student performance and progress as well as clarity of instructional goals had occurred since the implementation of accountability systems. Respondents suggested that grade advancement decisions were rarely made in the same way for students with and without disabilities.

Of the 249 participants, 155 participants responded to an open-ended question how grade advancement decisions were made for students with and without disabilities. Two of the authors involved in this study reviewed all of the participants' responses for common themes and determined that five main themes emerged. The five common themes were: a) general and special education students have different standards or requirements, b) special education students are rarely retained, c) students in special education are promoted with the understanding that services and support will be provided, d) grade advancements are made on a case-by-case basis by the IEP team, and e) no differences observed in grade promotion.

Thirty responses were randomly chosen in equal proportion from each category to ensure that the participants' responses could be categorized into main themes. Two

graduate students coded the responses across the six categories. Interrater agreement was 87.1%. Respondents were asked to rate the importance of 24 items in making grade advancement decisions on a scale of 1 to 4. The items represented four factors: student characteristics (e.g., effort, attendance, age), school performance (e.g., teacher-assigned grades and recommendations), test performance (e.g., performance on standardized and required tests), and other school indicators (e.g., the availability of intervention services, promotion policies). Each respondent's ratings were averaged over all of the variables within a factor. The results of the analysis indicated a significant main effect for respondent profession. More general education teachers than special education teachers reported student characteristics, school performance, and test performance as important in making grade advancement decisions.

The low response rate was a major limitation of the study. Another limitation was the procedures used for participant selection. Selection bias may have been a factor in the data since the principal of the schools chose who would receive the surveys. Although grade advancement decisions are not based on exit exam performances, in California exit exam performance does hinder a student's chance of receiving a high school diploma. Consequently, students who might be in danger of failing exit exams warrant further examination. The conclusions drawn from Christenson et al. (2007) implies that from the educators perspective, students with disabilities will have improved performance in school and on assessment measures as a result of the accountability system. Changes have been made to improved student outcomes, but no direct data were retrieved from students' actual performance.

Although Christenson's study seems promising, the fact still remains that students with disabilities fail at a much higher rate on standardized test than students without

disabilities. The study used a sample from elementary and middle or junior high school personnel which most likely has not had experience with the impact of high stakes assessment on students at the high school level where the stakes are higher. Therefore, the investigation of the relationship between placement for mathematics instruction (mainstream or self-contained) and performance on high-stakes tests at the high school level is needed. Given the effects of performance on the high stakes test, some researchers examined factors that would identify students most at risk of not passing.

Struggling students were the focus of an investigation conducted by Zau and Betts (2008) to identify students at risk of failing exit exams, like the CAHSEE, or the kind of interventions, if any, that might boost the achievement and hence increase the CAHSEE passage rate. The researchers' purpose was to identify students who would be at risk of failing the CAHSEE earlier in school to improve the students' chances of passing. The San Diego Unified School District (SDUSD) was used as a test case for this investigation. Zau and Betts (2008) focused on students in the class of 2006 which was the first class required to pass the CAHSEE to obtain a high school diploma.

The dataset included students' grades, test score on California Standards Test (CST), individual characteristics such as race/ethnicity, English learners (EL) and special education status, and characteristics of the students' school, peers, and teachers. Data were collected on how the cohort fared overall, both in the San Diego Unified School District and in California as a whole. The pass rates for seniors in 2006 and for these same students in 2004, at the end of grade 10, were collected. The pass rate margins—that is, how close some students who passed the exam actually were to failing the CAHSEE were also examined.

The investigation revealed the overall passage rates for the class of 2006 in San Diego and statewide was almost identical, at 91 and 90 percent, respectively. These pass rates did not include students who dropped out before 12<sup>th</sup> grade. Roughly one-quarter of the 10<sup>th</sup> grade students in 2004 failed to pass each component of the CAHSEE. Passage rates for special education subgroups were far lower than for the population at large. Of the 770 10<sup>th</sup> grade Students with Disabilities (SWD), not also classified as EL, who took the CAHSEE in 2003-2004 school year 177 (23%) passed. Of the 221 10<sup>th</sup> grade students classified as EL and SWD who took the CAHSEE, 4 (1.8%) passed compared to the 4,513 (67.4%) of the 6,697 Non-EL and Non-SWD 10<sup>th</sup> grade students who took and passed the test.

Zau and Betts (2008) used student data from grade 9 to predict grade 10 CAHSEE outcomes. The means, standard deviations, and ordinary least squares (OLS) models were used to analyze the data. The students who had passed the CAHSEE component by the given grade were used to analyze the data. African American students, English learners, and students with disabilities were significantly less likely to pass the CAHSEE in grade 10 (relative to white, non-EL, and non-special education students, respectively). The better a student performed in the classroom in terms of GPA and on the CST test, the more likely that student was to pass the CAHSEE. On the measures of the racial/ethnic composition of the student population at a school and the percentage of students eligible for free lunch did not appear statistically significant. Math test scores were a stronger predictor of success on the CAHSEE.

The researchers suggested that they could predict passage by 4<sup>th</sup> grade almost as well as they could by 9<sup>th</sup> grade. Math test scores on the CST in the elementary grades actually predict passage of the math portion of the exit exam better than the math scores

on the CST in the later grades. Academic grade point average (GPA) was the strongest predictor of eventual success or failure on the CAHSEE as well as students' behavior in the elementary grades. The study implies that educators can use information from earlier grades to predict whether students will be at risk for failing the CAHSEE to aid in early intervention strategies. A weakness of the study was that it did not fully explain factors, like setting, unique to the population of students with disabilities, an important population to the entire testing and accountability system.

Similar to Zau and Betts (2007), the current study used data from 10<sup>th</sup> grade students with disabilities at the high school level who are taking the CAHSEE for the first time. The current study expounded on the results of Zau and Betts (2007) by examining the relationship that setting for mathematics instruction had to the outcomes on the mathematics section of the CAHSEE for students with disabilities. The variables used in Zau and Betts (2007) was the same control variables used in the current study, so that any observed differences may be attributed to setting and not differences in ability. This was done because not all students with disabilities receive instruction in the same environment or have the same level of prior knowledge.

Schools today face parental, community, and political pressure to have students pass the CAHSEE and graduate from high school. Districts also face the requirement of meeting state and federal growth targets in relation to student proficiency levels in alignment with the federal NCLB Act, regulating that schools and districts make AYP and meet the states' API. As such, researchers have investigated factors which contribute to a failure to pass and/or receive a proficient score on the California High School Exit Exam in mathematics (Taylor, 2006). Specifically, the relationship between students'

eighth-grade mathematics courses and their results on the CAHSEE in terms of their passing and/or scoring at the proficient level has been explored.

Taylor (2006) used data from a suburban school district in Riverside County in Southern California. The data consisted of 2004 CAHSEE scores in mathematics of 3,114 students in the graduating class of 2006. The participants in the study attended one of four comprehensive high schools or one of the four alternative high schools in the school district. The sample consisted of 1,359 White and 1,323 were Hispanic students. The remaining 432 participants represented a variety of other ethnicities with no one ethnicity representing more than 6% of the sample. The school district's database system was used to obtain information on students' background information, demographics, academic performance, and achievement data. Information gathered included gender, ethnicity, socioeconomic status, parent education level, eighth grade math course level, eighth grade math letter grade earned, eighth grade scale score on the CST in mathematics, CAHSEE math scale scores, and CAHSEE pass/fail status.

The Statistical Package for the Social Sciences (SPSS) software was used to link CAHSEE results with academic, assessment, and demographic information. Correlations, descriptive statistics, multiple regressions, discriminate analysis or logistical regression were used to analyze the data. Taylor's research revealed that students' eighth grade mathematics course level had a strong relationship with their 10<sup>th</sup> grade CAHSEE math scale scores, CAHSEE proficiency status, and pass/fail results. While more than 91% of all students enrolled in eighth grade Algebra I or higher scored proficient on the mathematics portion of the CAHSEE administered in the 10<sup>th</sup> grade, less than 33% of all students enrolled in a course less than Algebra I in the eighth grade scored proficient on the 10<sup>th</sup> grade CAHSEE in mathematics. More than 99% of all students enrolled in eighth

grade Algebra I or higher successfully passed the mathematics portion of the CAHSEE administered in the 10<sup>th</sup> grade. However, 72% of all students enrolled in a course less than Algebra I passed the 10<sup>th</sup> grade CAHSEE in mathematics. Taylor (2006) concludes that 8<sup>th</sup> grade CST math scores and a student's math course letter grades were strong predictors of CAHSEE mathematics scores.

This study lends support to exposing students to more rigorous math curriculum at an earlier age although it is unclear whether this is also true for students with disabilities. The fact that students with disabilities were not addressed was a major limitation to this study. Students with disabilities are a mandatory element of the accountability system in California. Educators need to be able to identify which of the students with disabilities are at risk of failing the CAHSEE. The present study identified students with disabilities who are at risk for failure by examining the environment in which they receive mathematics instruction.

Nichols (2003) also conducted research to examine indicators that may predict failure on state high school graduation tests, but for performances in English/language arts and mathematics. Participants included students at six different high schools in a large urban school district. Data were compiled on students, with specific emphasis on the graduation classes of 2000 (n = 2,000), 2001 (n = 2,056), and 2002 (n = 2,364) who failed to pass the state of Indiana's minimum graduation proficiencies during their tenth grade year (1992-1999) in English/language arts or math, or both.

The data included background information on students who failed to meet the Indiana high school graduation requirements. The collection of data also included earlier test results from their tenth, eighth, sixth, and third grade years when these students were identified as being below state proficiencies at that time. Other information (e.g. history

of school attendance, noted by average yearly absences from school, and when appropriate, academic grade point averages for a number of years) was also collected. Raw data from the school district and student identification numbers to track students and their test scores were utilized.

The researchers used descriptive data collected by gender, ethnic background, and student economic status to analyze the data. A correlational matrix was used to analyze the relationship among yearly absences for the graduation class of 2000, 2001, and 2002 students who failed to meet state graduation proficiencies in English/language arts and math. Normal bell-curve equivalents for language, mathematics, and reading at the participants third, sixth, eighth and tenth grades on the state mandated standardized tests called the *Indiana Statewide Testing for Educational Progress (ISTEP)* was also used. More than 50% in each graduation class who failed to meet minimum competency standards on the statewide ISTEP graduation proficiency exam for English/language arts as sophomores were minority students, designated as having ethnic backgrounds of African American, Hispanic/Latino, Native American, and Asian American.

Students' standardized test scores varied inversely with their average yearly absences. As students absences increased, academic achievement decreased in each of the core subject areas. Results suggested GPA, absences or attendance by sixth grade, low SES, and students who struggled on standardized exams or failed to meet state minimum requirements in third grade were indicator of academic failure. Poor achievement in English was discovered as closely related to poor math achievement. The results lend support for GPA and prior academic achievement as predictors for passing the State of Indiana High School Graduation Exam.

A limitation to the study may be the transient nature of the student population. The authors noted the population, from 3<sup>rd</sup> grade to a student's senior year may fluctuate as much as 25-30%. The researchers did not state whether students with disabilities were included in the study. As such, this study specifically examined the students with disabilities population controlling for prior achievement to assess whether setting had an impact on pass rates.

Another variable which may have an impact on achievement in school is the thoughts and attitudes students with disabilities have of themselves. Research indicates that children with LD are often socially rejected by their peers and have difficulties establishing and maintaining friendships (Wiener, 2004; Wiener & Schneider, 2002) which may have an effect on their achievement. To develop a deeper understanding of students thoughts, feelings, attitudes --both positive and negative-- as well as the ways in which they present themselves Raskind, Margalit, and Higgins (2006) studied students' online messages. The descriptive study examined children's presentations of the learning disabled "(LD) experience" (p. 253) as expressed in online messages written on a pre-existing public website designed specifically for children with learning and attention problems. Raskind et al. stated children with LD are at greater risk for experiencing loneliness, low self-esteem, anxiety, and depression than non-disabled peers as mentioned in Margalit and Al-Yagon (2002). The research also explored children's descriptions of their difficulties and abilities. The researcher assumed that children viewed websites as a safe environment and they would disclose and share aspects of their LD identity and reveal their thoughts, feelings, and attitudes towards living with their LD freely.

The researchers hypothesized that disclosure regarding LD would encompass academic, social, and emotional domains. There was no protocol for selecting a sample of children with LD from an Internet website as one has not been previously established. The contact and collection of data directly from the research participants was not possible because the researchers indicated it would have threaten Internet anonymity and influenced the nature and authenticity of online messaging, as well as children's willingness to continue to use websites.

An online survey of 240 regular users of the website was administered in the fall of 2004. The survey was not part of the current research, but was used by the researchers to evaluate the impact of the site. The electronic survey also incorporated voice support was used to search for potential participants (i.e., the respondent could hear the question read aloud by a recorded human voice as an accommodation for children with reading difficulties). Two questions were included in the survey: "Do you have a learning disability?" and "Do you have difficulties with learning or paying attention?" Children who answered "yes" to the questions and "yes" to additional questions indicating specific difficulties in reading, writing, and/or learning were included in the study. Children who reported problems only in math were excluded.

Over 1,000,000 electronic communications (e-mails, message board submissions) from approximately 30,000 active, registered users were reviewed for content indicating the presence of LD. The sample consisted of 164 children, ages 9 to 18 with a mean age of 12.5 who were "self-identified" as LD. Participants were drawn from registered users of a free public website throughout the U.S. Of the 164 children participating in the study, 108 were female and 56 were male. No information was available regarding the

participants' SES or ethnicity. Additionally, no diagnostic information was available. Children accessed the website from home and school.

SparkTop.org, the website used to study the online communication of children with LD, was developed by Schwab Learning, a program of the Charles and Helen Schwab Foundation. Although open to all children, the site was designed for children with learning and attention problems, ages 8-12, to provide an online experience and create a virtual community where children could "connect" with other children, build self-esteem, develop self-awareness, enhance self-advocacy, gain knowledge of learning strategies, create art, play games, as well as receive accurate information about learning and attention problems.

The website used current technological tools, design principles, and "pop" culture trends in children's media and entertainment to ensure the creation of a "cool site," and not a segregated site for "children with problems." To use the site, children must register through an adult (parent, guardian, or a teacher) by means of Privo, a parent permission service. No personally identifiable information was put on the site (e.g., name, address, phone number, e-mail address) and website staff. The authors considered the site to be a safe and secure base for children with LD to self-disclose and share their thoughts, feelings, and experiences of living with a LD.

The research team consisted of three individuals. Each researcher had over 30 years of teaching, clinical, and research experience in the field of LD. Four thousand nine hundred and three e-mails sent from the 164 self-identified LD participants to other users on the site between 2003 and 2005 were reviewed. The data from the website were converted into a Microsoft Excel file to facilitate data management and analysis. The

database contained information on the age and sex of both sender and recipient, date/time of messages, and the actual written message.

A content analysis was conducted to search for expressions of "LD self-presentation" in the online messages and examine children's communicative behavior. The data analysis searched for themes that represented participants' experiences and portrayals of their LD. Each researcher developed and wrote his or her own codes or category labels directly on the hard copies next to specific messages. In some instances codes represented general themes (e.g., academic difficulties, social problems) while some codes were more specific and reflective of general theme subcategories (e.g., reading problems, teasing). Independent analysis alternated with regular meetings between the researchers occurred over a 12- month period. E-mails were sorted according to agreed upon thematic categories. Messages ranged from 1 word to 202 words, with an average length of 22 words. Emphasis was placed on capturing and reporting the "insider's perspective" from the informants "living" within this virtual community. Direct quotations from the participants were used to verify and elucidate identified themes, as well as represent the shared experience of living with a LD among the members of this virtual community.

Six major themes emerged from the analysis of children's messages: LD identity, disclosure of academic difficulties, disclosure of emotional attitudes, disclosure of social/interpersonal issues, asking for help, and positive LD. Many identified themselves as individuals who belonged to a group of children with LD, communicating on a site for children with LD. Having LD was considered a problem as well as an identity. The children presented their "overall identity" by providing information about their age, sex, and their location in terms of a state or city. Several children provided physical

descriptions, disclosed information about their families, revealed personal preferences regarding youth culture, and discussed hobbies. Many messages contained information about the children and their areas of competence and interest.

Individual differences were reflected in the children's insights into the types, severity, and pervasiveness of their difficulties. Additionally, individual differences were expressed in their emotional tone with several children expressing clear distress, while others appeared accepting of their difficulties as part of their identity. Although expressions varied, the majority of children who wrote messages disclosed a negative emotional attitude towards their LD and specifically towards their academic difficulties. In some cases, the children's expressions noted that having LD affects their self-worth. The messages revealed expressions of emotional distress, including feelings of sadness, diminished self-worth, loneliness, and fear, which resulted from their academic struggles, social rejection, and family stress.

The children also reported personal and social difficulties as well and viewed their social exclusion as part of their LD identity. They expressed that having a LD has social ramifications (e.g., teasing, ridicule, social rejection) and expressed a fear of sharing their LD identity with friends. The children expressed hypothetical fears, as well as difficult and hurtful social experiences. The emotional expressions of social distress were often accompanied by expressions of dissatisfaction with the LD identity, even a readiness to give up accommodations necessary for school success. The children expressed a desire to be treated similarly to their peers in order to avoid being identified as different, rejected, and even abused by others. A vast majority of children asked for help. The children complained that they were distressed and not getting the help they needed.

Approximately two-thirds of the requests for help and advice from the website users were directed towards the teen mentors and the adult "LD Expert" however, children also routinely reacted to one another, sharing worries and frustrations. Only a few children made positive statements regarding their LD. Due to legal, practical, and ethical concerns, the researchers were not able to confirm the identification of the sample as LD. The study only included children who self-identified as LD. Thus, it is unclear whether the intended population, LD, was accurate. However, this descriptive study provided some insight into the experiences of students with disabilities, their thoughts about being learning disabled and suggested that students with disabilities struggle with academics and more. Raskind et al. (2006) examined the academic, social, and emotional difficulties faced by students with disabilities which can result from negative societal experiences, whether directly or indirectly, related to the identification of a student with a disability. The current study expanded on the idea of group identification and its relation to academic performance through the examination of the relationship of stereotype threat and educational setting for students with disabilities.

The current section reviewed literature relative to high stakes testing and performance of students with disabilities. Specifically the section examined literature related to the performance of students with disabilities on high stakes end-of-course tests. Within this section was the review of the effects of high stakes assessment for students with disabilities and whether differences existed between students with disabilities and students without disabilities. The literature reviewed addressed student factors and their relationship to performance on high school exit exams which would assist in the identification of students at risk for failure on the high stakes tests. The section ends with a review of literature of students' with disabilities experiences and their self-perceptions.

The literature indicated that students' with disabilities performance is lower than students without disabilities. The pass rates in an academic Algebra 1 course for students with disabilities have increased over time, but the increase is slow. This is important because Algebra 1 standards are a part of the CAHSEE mathematics section. Moreover, high stakes assessments have increased instructional changes and increased monitoring of student progress and performance to improve student outcomes. Students with disabilities indicated their belonging to the group of children with a learning disability and the difficulties, academic, social, and emotional relative to the group identity.

### **Stereotype Threat**

Stereotype threat theory has been used to explain underperformance of minority students in academic domains and of women in math performance. The role of anxiety as the mechanism (or a major aspect of the mechanism) through which stereotype threat affects academic outcomes has been of growing interest. For example, Osborne (2001) examined whether anxiety could explain the racial differences in academic performance between White students and students of color (e.g., African Americans, Latinos, and Native Americans) and gender differences in math performance.

Two hypotheses were tested. The first hypothesis, racial differences in academic outcomes (specifically, achievement test scores) were at least partially explained by differing levels of anxiety while engaged in academic evaluation or while in the academic environment was tested. The second hypothesis, gender differences in math performance (math achievement test scores) were at least partially mediated by differential levels of anxiety was tested.

Data for the study were drawn from the senior cohort data file from the High School and Beyond (HS&B) study initiated in 1980 by the National Center for Educational Statistics (NCES). The senior cohort consisted of 28,240 seniors from 1,015 schools in the United States. Sex, race, and ethnic background variables were gathered during the same data collection session as the other variables. Five racial groups were formed from the combination of race and ethnic background variables. The five racial groups were White ( $n = 12,557$ ), African American ( $n = 1,846$ ), Latino ( $n = 1,047$ ), and Native American ( $n = 111$ ). Three separate analyses were conducted to address whether anxiety explained racial differences in academic outcomes, one for each of the White vs. non-White contrasts. For each White/Non-White contrast analyses, the White group was randomly sampled so each group contained equal cases. Full replacement was used in the White group for each sampling.

Students in the sample completed standardized achievement tests covering vocabulary, reading, and mathematics. All sections of the standardized tests were timed. The vocabulary test consisted of 27 synonym-format items. The reading test consisted of 20 items related to several short (100–200 word) passages and the mathematics test consisted of 33 items in a format where students were asked to indicate which of two quantities is greater. Immediately following the standardized test the subjects completed a short battery of questions used to assess anxiety. All questions on the anxiety measure used a dichotomous *yes* (1) and *no* (0) scale with the item stem “How did you feel while you were taking the tests?” and included the following items: tense, under pressure, under strain, nervous/jittery, uneasy, calm, afraid of not doing well, and uncomfortable. All items were scored so that higher scores indicated more anxiety. Prior academic preparedness was assessed using student reported grades in high school.

The standardized scores on the three tests were correlated. As such, a composite achievement test score was created by averaging the three standardized scores. A composite anxiety score was created by summing the variables. The composite variable had significant positive skew. To reduce skew, all scores over 6 (less than 1.1%) were recoded to 6. Additionally, a quadratic anxiety term was created by squaring the variable.

Three univariate ANOVAs were performed, each contrasting Whites with African Americans, Latinos, or Native Americans to address the relationship of race to achievement and race to anxiety. To test whether anxiety related to achievement test, a multiple regression was computed predicting achievement from anxiety and a quadratic anxiety score. Two multiple-regression analyses were performed to test mediation for each racial contrast. The first included only the race variable predicting achievement test scores. The second was a blockwise multiple-regression analysis that entered the anxiety and a squared anxiety term (as the anxiety and performance literature had reported curvilinear relationships between the two variables) on the first step and the dummy-coded race variable on the second step. To test whether the relationship between race and achievement was significantly reduced, the difference between the two regression coefficients for race was tested for significance. Results indicated that White students performed significantly higher than any of the non-White groups when comparing race to achievement and White students had lower anxiety scores than the non-White students. While anxiety did not completely mediate the relationship between race and achievement, the race effects on achievement were partially mediated in two of the three analyses (White vs. African American and White vs. Latino).

Since anxiety was measured after completing the last item of the test, there was a question of reciprocity, meaning whether the differences in anxiety were caused by

differences in test results. To address this issue two additional analyses were performed: one replicating the previous analyses controlling for academic preparation and one testing whether there are significant differences in anxiety once test scores are covaried (i.e., to test whether differences in anxiety are wholly attributable to differences in test scores). Results indicated after covarying prior academic preparedness, anxiety significantly mediated the relationship between race and achievement for the same two analyses (White vs. African American and White vs. Latino). Race accounted for 10 to 15% of the variance in achievement test scores, and anxiety accounted for 28 to 31% of the effect.

Three separate ANOVAs were calculated, each contrasting Whites with African Americans, Latinos, or Native Americans to test whether there were significant racial differences in anxiety once achievement was controlled. Again, results suggested significant differences in anxiety for both White vs. African American and White vs. Latino contrasts, but not for White vs. Native American. The results indicated African American and Latino students were more anxious than White students.

To address the second hypothesis, all students with valid data were retained (10,320 males and 11,510 females). Females were randomly sampled to equalize cell sizes at 10,320. ANOVAs were performed to analyze the relationship of sex to achievement and sex to anxiety. A multiple regression was used to predict math achievement scores from anxiety and its quadratic component. To examine whether the introduction of the mediators into an analysis would significantly attenuate the relationship between sex and math achievement test scores, two multiple regression analyses were performed. The first included only sex to predict math achievement test scores. The second included anxiety and its quadratic component on the first step and sex on the second step. The results supported a partial mediation explanation. Anxiety (and

its quadratic component) accounted for a significant portion of the achievement test differences between Whites and African Americans and between Whites and Latinos. Anxiety was not found to explain the differences between White and Native American Students, which may have been attributed to the smaller sample size for this group.

Since anxiety was measured after the test was completed, it is not entirely clear that anxiety caused differences in test performance. The study addressed this weakness by assessing anxiety along with other moderating factors. It was unclear to what extent students were able to perceive their own performance accurately as no feedback was given to them regarding their performance. With the CAHSEE students will receive feedback regarding their performance on the test within approximately 6-8 weeks.

Osborne's (2001) study examined high school seniors which developmentally was similar to other studies related to stereotype threat theory. In the current study, students in their sophomore year of high school were examined as this was the first administration of the high-stakes exit examination for graduation with a high school diploma that the student was able to take.

Much of the focus on testing is the result of the No Child Left Behind Act of 2001 (NCLB, 2002), which has forced states to implement statewide accountability systems. Furthermore, NCLB requires that the test results be subdivided by poverty, race, ethnicity, disability, and limited English proficiency to ensure that these groups will not be left behind (U.S. Department of Education, 2001). This subdivision of test scores places a focus on the differences between groups of students. As such, Kellow and Jones (2008) examined the extent to which African American high school freshman students experience stereotype threat when taking a test that is seen as a predictor of their success on a high-stakes test. The widely held negative stereotypes related to the intellectual

performance of African Americans and the increased reliance on standardized tests as a means to measure student progress created a set of conditions that would appear conducive to African Americans experiencing stereotype threat during testing.

Participants were recruited from freshman introductory algebra courses at a large urban high school in Florida. Informed consent forms for parent or guardian approval were provided to 641 students. The study was presented as an investigation of mathematical reasoning. For their participation, students received a gift card for a free meal at a local restaurant. One hundred eighteen (18.4%) of the forms were signed and returned. The ethnic composition for students returning forms was White 53%, African American 42%, Hispanic 3%, and Asian 2%. Over half (66%) of the students were female. On the day of testing, 11 students (8 African American and 3 White) were absent and 1 student failed to complete the dependent measure. Since the focus of the study was on African American and White students and because so few Hispanic and Asian students participated (combined  $n = 5$ ), data for the Hispanic and Asian participants were excluded from the final sample. As a result, the total number of student participants was 101. All participating students were average to excellent in mathematics performance based on the normal curve equivalent (NCE) scores obtained by each group on their eighth-grade standardized tests.

The researchers used the school's database to classify the participants' ethnicity and gender. Students were stratified according to ethnicity and then randomly assigned to either the experimental (evaluative) or comparison (non-evaluative) condition. The evaluative condition referred to the group of students who were told that their test scores would predict how well they would perform on a standardized high-stakes test. The non-evaluative condition referred to the students who were told that the test they were about

to take was unbiased. A total of 48 students were in the evaluative condition group and 53 students in the non-evaluative condition group. The testing took place during the students' regular mathematics class period.

After arriving at one of two classrooms reserved for the experimenters (who were both White males), participants were provided with a pencil and a booklet containing the student assent form, the *Applied Personnel Research Spatial Ability Test* (APR; Wiesen, 1996), and a questionnaire. After reading and signing the consent form, students received instructions, one for the evaluative group and one for the non-evaluative group. Following the instructions, students were guided through an example item to ensure that they understood the proper procedure for interpreting and responding to the items. Next, students completed the item measuring their proximal expectancy for success.

Students were given five minutes to work on the APR, then were told to stop (whether they had attempted all of the items or not) and complete the questions on the following pages, which contained the achievement goal orientation items, the *State Trait Anxiety Inventory* (STAI; Spielberger, Gorsuch, & Lushene, 1970), the *Self- and Task-Perception Questionnaire* (STPQ; Eccles & Wigfield, 1995), and the stereotype threat-specific scale. When students finished, they handed in their answers and returned to their mathematics class. Soon after the experiment, the researchers revisited the school and debriefed students as to the true purpose of the study as well as answered any questions the students had pertaining to the experiment. In addition, students were provided with a debriefing sheet to share with their parent or guardian.

The APR served as the primary dependent variable in the study. The APR is a personnel selection test for skilled clerical applicants. It is a timed test consisting of 50 items that relate to a set of blocks that are stacked in various configurations. Respondents

must indicate how many other blocks are touched by a specific block. This type of problem was consistent with the types of knowledge and skills that were expected of ninth-grade students. Prior to attempting the APR, students' proximal expectancy for success using a single item was measured. Students responded on a 7-point Likert-format scale anchored at 1 (*not well at all*) to 7 (*very well*). Adapted items were used to assess students' goal orientation during the test. One item was adapted to measure students' performance-avoidance goal orientation and another was adapted to measure students' performance-approach goal orientation.

Students responded on a Likert-format scale anchored at 1 (*not at all true of me*) to 7 (*very true of me*). The *STPQ* has been used with adolescents to assess a variety of constructs related to their beliefs, values, and attitudes regarding an academic domain, such as mathematics. The five-item scale that pertained to ability perceptions and expectancies for success was used. Each item was scaled using a 7-point Likert-format scale anchored at 1 (*much worse*) to 7 (*much better*). A short-form of the *STAI* consisting of eight items with a 4-point Likert-format scale anchored at 1 (*not at all*) to 4 (*very much*) was used to assess state anxiety of students while taking the APR. Items were summed to produce a total score. The extent to which students expected to be stereotyped on the APR based on their ethnicity was measured using a 5-item Likert-format scale anchored at 1 (*strongly disagree*) to 5 (*strongly agree*).

An unbalanced  $2 \times 2$  ANOVA using no covariates was employed. All analyses were conducted with SPSS using the General Linear Model procedure with adjusted Type III sums of squares to compensate for the unbalanced nature of the design. On the APR, participants attempted an average of 27.8 items out of the 50 possible items. However, they only successfully completed an average of 16.7 items for an average of 60.1%

correct on the attempted items. Hypothesis 1 was that African American participants in the evaluative condition would score lower on the APR than White participants in the same condition. While the interaction effect was statistically significant, African American students in the evaluative condition performed almost identically to African American students in the non-evaluative condition. With respect to White students, however, there was a sharp difference between conditions, with students in the evaluative condition outperforming students in the non-evaluative condition.

A second hypothesis predicted that African American students in the evaluative condition would report lower self-perceptions of ability and expectations for success (as measured by the STPQ) than White students in the same condition. In the evaluative condition, the mean of African American participants on the STPQ was slightly higher than that of White participants. Similarly, African American students in the non-evaluative conditions had a slightly higher mean than White students. None of these differences, however, were statistically significant. A third hypothesis was similar to the second, except that instead of measuring students' perceptions of ability and expectancies for success in the domain of mathematics, students' proximal expectancy for success specifically on the APR that they were about to complete was measured.

African American and White students in the non-evaluative condition had nearly identical proximal expectancy for success on the APR. However, in the evaluative condition, African American participants had a statistically significantly lower mean proximal expectancy for success than White participants. Regarding achievement goal orientation, it was predicted that African American students in the evaluative condition would have higher performance-avoidance goals and lower performance-approach goals than White students. However, there were no statistically significant effects associated

with either dependent variable. With respect to performance avoidance, African American students in the evaluative condition were very similar to White students in their responses. Similarly, African American students in the non-evaluative condition were very close to White students in terms of mean differences.

The data related to performance approach mirrored those of the performance-avoidance variable. Although it was hypothesized that African American students in the evaluative condition would experience more state anxiety than White students in the same condition, statistically insignificant differences between condition and ethnicity were found. The researcher's measure of stereotype threat-specific found a statistically significant and substantive main effect. African American students' scores in both the evaluative and non-evaluative conditions were higher than those of White students in both conditions. White students scored statistically significantly higher than African American students when told that their test performance would be predictive of their performance on a statewide, high-stakes standardized test. This finding indicates that African American students in both conditions were more likely than White students to believe that they were being stereotyped by others. A lack of degradation in performance of African American students in the evaluative condition and the profound performance increase of White students in the same condition was revealed, despite the fact that African American students had significantly lower proximal expectancies for success in the evaluative condition than in the non-evaluative condition. The researchers concluded the proximal measure of expectancy for success at the task level was a more sensitive measure than a domain-level measure of abilities and expectancies.

Kellow and Jones' (2008) study suggests that freshman African American high school students were aware of the negative stereotypes related to statewide standardized

tests and that these stereotypes have an effect on the expectation of success on a particular test. African American participants in the evaluative condition did not report higher levels on the stereotype threat-specific measure than African American participants in the non-evaluative condition as was predicted. The authors suggested this resulted from other variables that were not measured in this study that account for the discrepancy, such as student characteristics or the school environment. A limitation to the study was the absence of high stakes. The students knew that the test was a part of the research and that no consequences would be attached for low performance, which would have had an impact of the results of the study. In addition, the pressure of confirming a negative stereotype would also be diminished as a result of the reduced threat of the high stakes assessment.

The Kellow and Jones' (2008) study is relevant to this study because both studies investigated the threat of stereotype and its relationship to student performance. Although the instrument used in Kellow and Jones (2008) may have been seen by the participants as a reduction in pressure to perform, the current study used a high stakes test, CAHSEE, as the measure of performance. Additionally, student characteristics (disability) and school environment (mainstream or self-contained) was accounted for. Kellow and Jones (2008) used perception of ability, goal orientation, and anxiety as mediators of stereotype threat; however, other researchers have investigated locus of control as a moderator of stereotype threat.

Previous research has shown that individuals with internal locus of control tend to be more competent and more highly motivated to perform well in academic tasks than individuals with external locus of control. Individuals who were competent and who cared about their performance were most vulnerable to the effects of stereotype threat

(Steele, Spencer, & Aronson, 2002). Therefore, individuals with internal locus of control appear to be candidates for the effects of stereotype threat. Cadinu, Maass, Lombardo, and Frigerio (2006) tested the hypothesis that individuals with internal locus of control will show a decrease in performance in a stereotype threat versus individuals with external locus of control who are not expected to show this effect due to a no-threat condition.

Participants included 154 students (80 female and 74 male) in their last year at a scientific high school in the northeast of Italy. In the Italian high school system, the scientific high school offers the strongest background in mathematics and is aimed at preparing students who will pursue university majors in scientific disciplines. In the Italian high school system classes were composed of 20–25 students who spend the entire school day together. Participants took part in the study on a voluntary basis. Participants attending the last year of high school were given a multiple-choice test that could be interpreted as measuring either logical or social intelligence.

A pretest, using a separate sample, was conducted to test whether logical intelligence was considered more typical of men than women and social intelligence more typical of women than men. An independent group of 57 students at the University of Padova (32 females and 25 males) were presented with the definition of logical or social intelligence and were asked to rate the degree to which that type of intelligence was considered feminine or masculine in society. The results suggested logical intelligence as a typically male characteristic, and social intelligence as a typically female characteristic.

For the study, a female experimenter handed out the questionnaires in the presence of one of the teachers, who remained in the classroom for the entire duration of

the experiment. The questionnaire included first the *Locus of Control Scale* (Rotter, 1966), followed by the experimental manipulation and then the performance test. Participants were randomly assigned to one of the two experimental conditions or to the control condition. In the two experimental conditions, written instructions informed participants that studies in social psychology have demonstrated that people with high levels of logical intelligence (or social intelligence, depending on the condition) have more success in life. Participants were then told that the goal of the test was to investigate whether there were differences between men and women in the degree to which they possessed and used logical (or social) intelligence. Participants in the control condition were simply told that a test would follow. The same test was used in all three conditions but labeled differently.

After the test, participants were asked to give an estimate of their performance by responding to several questions. Two predictions were examined in the study. First, it was predicted female students would perform worse in the logical intelligence condition (stereotype threat for women) than in the social intelligence condition (no-threat condition for women) whereas male students would perform worse in the social intelligence condition (stereotype threat for men) than in the logical intelligence condition (no-threat condition for men).

To ensure that performance in the different experimental conditions was a function of the different test labels (social intelligence vs. logical intelligence) rather than being the result of how much importance participants attributed to the different types of intelligence, participants in the logical intelligence condition were asked to rate how important it is in life to have a good level of logical intelligence and participants in the social intelligence condition were asked the same question regarding social intelligence.

No such question was presented to control participants. Responses were rated on a scale ranging from 1 (*of little importance*) to 7 (*very important*). Second, it was predicted that internal locus of control individuals would show a stronger decrease in performance in the stereotype threat condition as compared to external locus of control individuals. The locus of control scale was administered before the intelligence test as a way to test for the potential role of locus of control in moderating the vulnerability to stereotype threat.

In measuring locus of control, a 12-item reduced version of the *Rotter's Locus of Control Scale* was administered. Participants were divided into two groups based on the median split of the locus of control distribution. Participants with scores lower than 0.51 ( $n = 72$ ) were named external locus of control individuals and participants with scores higher than 0.51 ( $n = 82$ ) were named internal locus of control individuals. The *Logical (Social) Intelligence Test* was similar to those employed in many Italian Universities for admission selection. The test was composed of 7 multiple-choice questions that could be interpreted as either logical or social dilemmas. The average percentage of correct responses was 47% and only 6% of participants responded correctly to at least 70% of the items.

Several analyses were performed in the study. For the pre-test, a 2 x 2 between subjects ANOVA was performed on the gender typicality ratings. The pre-test scale ranged from 1 (*typically feminine characteristic*) to 7 (*typically masculine characteristic*). A test of whether the mean scores for logical intelligence and social intelligence were significantly different from the scale midpoint of 4 was done because the midpoint was considered neutral (neither masculine nor feminine). This was done using a one-sample t-test with the scale midpoint of 4 as the test value. A 2 x 2 ANOVA was performed on the importance attributed to the ability that had been measured (logical

or social intelligence) with gender and type of intelligence as between-participants factors to test the importance of logical vs. social intelligence. Additionally, a 2 x 3 ANOVA was conducted on the sum of correct answers with gender (female vs. male) as the first between-participants factor and type of test (Logical, Control, Social) as the second between participants factor to investigate the effect of test labels on performance. Type III sums-of-squares were used to correct for non-orthogonality present with unequal frequencies across the cells.

To investigate the moderating role of locus of control, participants' data were aggregated such that the Social Intelligence condition was recoded as threat for the male participants and no-threat for the female participants. The logical intelligence condition was recoded as threat for the female participants and no-threat for the male participants due to a small number of participants in the cells of a three-way interaction that would have been used to test their hypothesis. The hypothesis was tested using a 2-way interaction between locus of control (internal vs. external) and threat condition (threat vs. no-threat). Since almost all control participants perceived the test as a logical intelligence test and showed exactly the same results as the logical intelligence condition, the control participants were not included. A 2 x 2 x 2 ANOVA was conducted on the performance score with Condition (threat vs. no threat), Locus (internal vs. external) and Gender (male vs. female) as between-participants factors. Tests for simple effects were performed to examine the effect of condition separately for internal locus of control and external locus of control participants. In addition, a separate analysis was conducted to include only participants in the control condition. Results of the pre-test indicated a main effect of type of intelligence was found with logical intelligence showing a significantly higher mean than social intelligence. The ratings of both logical intelligence and social intelligence

were statistically different from the scale's neutral point. No gender effect or interaction emerged.

Twenty-three participants from the experimental conditions, 20 in the Social Condition and 3 in the Logical Condition were excluded from the manipulation check because they chose a different type of intelligence than the one suggested in the test description (i.e. they did not seem to believe the experimental manipulation). Since most of the participants (50 out of 53) considered the test as a logical intelligence test, the three participants who considered the test either a social intelligence or a creativity test were excluded from the analyses in order to maintain homogeneity in the control sample. However, it was noted that the same results were found when all participants, regardless of their responses on the manipulation check, were included in the analyses. The researchers concluded there was not an effect for importance of logical versus social intelligences.

When the effect of test labels on performance was examined, a main effect of gender was found with male students obtaining higher scores than female students. A highly significant interaction between gender and type of test was found. Female participants obtained higher scores in the social intelligence condition compared to the logical intelligence and control conditions. To the contrary, male participants obtained lower scores in the social intelligence condition than in the logical intelligence and control conditions. No differences were found between the female logical intelligence, the female control, and male social intelligence conditions. Furthermore, no differences were found between the male logical intelligence, the male control, and the female social intelligence conditions. The estimation of one's own performance showed a main effect

of gender was found with males reporting higher estimates than females. No other effects were found.

A main effect of condition was found for moderating role of locus of control and a main effect of locus of control was found with internal locus participants performing better than external locus participants. The interaction between condition and locus of control was significant. The effect of condition was significant for the internal locus of control participants but not for the external locus of control participants. No gender effects were found. A separate analysis included only participants in the control condition resulted in internal locus of control participants performing better than external locus of control. Although when comparing threat and no-threat conditions, internal locus of control individuals showed a sharp decrease in performance whereas external locus of control individuals showed no significant decline in performance as a result of stereotype threat.

Equally important to the participants was logical and social intelligence. The stereotype-related deficit did not seem to be the result of these individuals complying with the expectations of their group's stereotype. To the contrary, internal locus of control individuals are typically over-achievers who seem to be fighting against the possibility of failing rather than resigning themselves to what is expected of them on the basis of the group stereotype. Despite their effort, these individuals seem to decline in performance under the stereotype threat pressure while other individuals (with external locus of control beliefs) remain unaffected.

The results of Cadinu et al. (2006) confirm the idea that internal locus individuals are 'over-trying' rather than 'under-trying' which affected performance. The implication for this study suggested that due to the high stakes nature of the CAHSEE, one's group

performance relative to his or her abilities will be depressed because he or she is fighting against the possibility of failing the test. This may confirm a negative stereotype about the group while another group's performance will be unaffected. The current study applied this phenomenon to students with disabilities in different group settings.

Cadinu, Maass, Lombardo, and Frigerio (2006) noted some participants did not believe the test label manipulation. Thus a second experiment aimed to replicate the results of the first experiment on the moderating role of locus of control in stereotype threat performance deficits was conducted. The goal in the second experiment was to strengthen the generalizability of the results from the first experiment using a different research paradigm in which women were confronted with an explicit threat regarding their gender abilities in math.

Experiment 2 differed from Experiment 1 in three ways. First, the locus of control scale was administered one week before the logical-mathematical test. Second, an explicit manipulation of stereotype threat with only two experimental conditions: a stereotype threat condition, reporting differences between men and women in logical-mathematical tests, and a no-threat condition, in which women were reassured that there were no differences between men and women in logical-mathematical tests was used. Third, only females were included in the sample to directly test whether the effect of an explicit stereotype threat in the logic-math domain would be moderated by participants' locus of control beliefs.

Sixty female psychology students from the University of Padova participated in the second experiment. Participants were contacted in class by a female experimenter and asked if they were interested in participating in a psychology experiment. All students were asked to fill out a brief pre-test questionnaire containing the locus of control scale.

Male students were also included in the pretest to prevent making gender salient during the beginning phase of the experiment. The experiment took place the following week with sessions of 3–5 female participants who performed the task individually. The experimenter was the same person who administered the pre-test a week earlier. Adapting a design originally proposed by Spencer, Steele, and Quinn (1999), a stereotype threat (gender-difference) condition and a control (no-gender-difference) condition were included for the experimental manipulation. Participants were divided into two groups based on the median split of the locus of control distribution. Participants with scores lower than 0.63 ( $n = 34$ ) were named external locus of control individuals and participants with scores higher than 0.63 ( $n = 25$ ) were named internal locus of control individuals.

The material for the experiment consisted of a difficult mathematics test composed of seven exercises (average number of correct responses = 4.4;  $SD = 1.77$ ). A  $2 \times 2$  ANOVA was conducted on the performance score with condition (threat vs. no-threat) and locus of control (internal vs. external) as the between participants factors. Participants in the stereotype threat condition performed worse than participants in the no-treat condition. A significant interaction between condition and locus of control was found. The internal locus participants performed worse under stereotype threat than in the control condition. By contrast, external locus participants performed similarly under stereotype threat and in the control condition. Although the main effect of locus of control was not significant, results in the control condition were in line with results from experiment 1. Results from the stereotype threat condition suggested that individuals with internal locus of control are at greater risk for performance deficits when exposed to stereotype threat. The implications for the current study was that students with disabilities

who are preoccupied with disconfirming a negative stereotype about students with disabilities will have lower performances on high stakes assessment than the students with disability for whom group identification is not a concern.

The current section reviewed literature relative to stereotype threat as an expansion of underperformance of stereotyped group identification. Research discussed in the section related to anxiety, goal orientation, performance perception and expectation, and locus of control as mechanism for which stereotype threat affects students' academic outcomes. The results of this review suggest anxiety as a mechanism for which stereotype threat is activated. In addition, this review supports the hypothesis that stereotype threat has a negative effect on student performance.

### **Summary of the Review of the Literature**

Several factors were presented in this chapter that may correlate with the rate of passage or failure on the California High School Exit Exam mathematics section for high school sophomore students with disabilities taking the exam for the first time. These factors were contained within the three categories of mainstreaming verses self-contained, high stakes test and students with disabilities, and stereotype threat. Combined, these factors may contribute to the poor academic success on standardized tests for students with disabilities at the high school level.

The first section reviewed literature focused on the differences in academic instruction for students with disabilities. This review indicates that for math instruction, students with disabilities placed in mainstream settings have better academic outcomes than students with disabilities in self-contained settings. Several limitations were noted in section one. First the studies lacked the ample sample size needed to accurately

generalize the findings. A second limitation was that the majority of the participants were from grades lower for which the stakes of performance is lower.

In section two, studies relative to the high stakes testing environment and students with disabilities were reviewed. This section revealed that performance of students with disabilities continues to be lower than students without disabilities. Moreover students who have disabilities often experience social, emotional, as well as academic challenges which have an effect on their performance.

The last section reviewed literature relative to stereotype threat. Stereotype threat is a phenomenon that has been used to explain student underperformance in standardized testing situations. The literature in this section supports the phenomenon. A limitation to the section regarding stereotype threat is that it has not been applied to the population of students with disabilities.

In summary, the ability of secondary students with disabilities to receive a passing score on the mathematics section of the high school exit exam is presumed to be limited by the setting in which math instruction is received. Moreover, the threat of confirming a negative stereotype about one's self (Steele & Aronson, 1995) diminishes student performance in the standardized testing environment. Since a high school diploma requires the passage of a high school exit exam, the extent to which secondary high school students with disabilities will receive their diploma will depend on their ability to receive a passing score on the exam. As such, an investigation involving high stakes testing, pass rates, and setting for students with disabilities is need.

## **CHAPTER THREE**

### **METHODOLOGY**

This chapter is organized into several sections. An introduction to the primary purpose of the study is presented first, followed by a description of the methodological research design employed in the study. The third section provides a description of the participants and their respective schools followed by a discussion on the protection of human subjects. The fifth section presents the procedures used in the development of the questionnaires including results from a validity panel study. Next, findings from the questionnaire pilot study are discussed followed by the procedures for administering the questionnaire and conducting the review of students' educational records. This chapter closes with a summary and a brief prologue to Chapter IV.

The primary purpose of this study was to examine students' anxiety associated with stereotype threat in relation to the California High School Exit Exam (CAHSEE) scores in mathematics for student with disabilities in mainstream or self-contained classroom settings for math instruction. This study addressed the following research questions:

1. When controlling for prior math ability (as measured by Algebra 1 grade, 8<sup>th</sup> and 9<sup>th</sup> grade CST scores on Math), is there a significant mean difference in the scores on the mathematics section of the CAHSEE for 10<sup>th</sup> grade students with disabilities enrolled in mainstream or self-contained settings for math instruction?
2. When controlling for prior math ability (as measured by Algebra 1 grade, 8<sup>th</sup> and 9<sup>th</sup> grade CST scores on Math), is there a significant mean difference in stereotype threat between 10<sup>th</sup> grade students with disabilities enrolled in mainstream or self-contained settings for math instruction?

3. When controlling for prior math ability (as measured by Algebra 1 grade, 8<sup>th</sup> and 9<sup>th</sup> grade CST scores on Math), is there a significant mean difference in anxiety between 10<sup>th</sup> grade students with disabilities enrolled in mainstream or self-contained settings for math instruction?

### **Research Design**

The study employed a causal comparative design used to compare two settings (mainstream and self-contained) with respect to three variables (CAHSEE, anxiety, and stereotype threat). The research investigation consisted of quantitative data collection that involved the administration of two questionnaires followed by a review of the students' educational records and their results on the math section of the CAHSEE. Students' prior math ability as measured by the students' California Standards Test (CST) math scores (8<sup>th</sup> and 9<sup>th</sup> grade) and their Algebra 1 course grade served as a covariate.

The research study included students who were classified as having a disability, enrolled in the 10<sup>th</sup> grade, taking the CAHSEE for the first time, and enrolled in a public high school in northern California. Since existing groups (mainstream and self-contained) in the educational settings were utilized, a causal comparative design was used. This method provided a process of examining information on the problem of high first-time failure rates on the CAHSEE for students with disabilities, and allowed for both descriptive and inferential data analysis in reporting the findings (Creswell, 2005).

Figure 1 below illustrates the hypothesis that the presentation of a high stakes test used for graduation from high school triggers anxiety and stereotype threat which

interferes with the achievement of students with disabilities in mainstream or self-contained settings for math instruction.

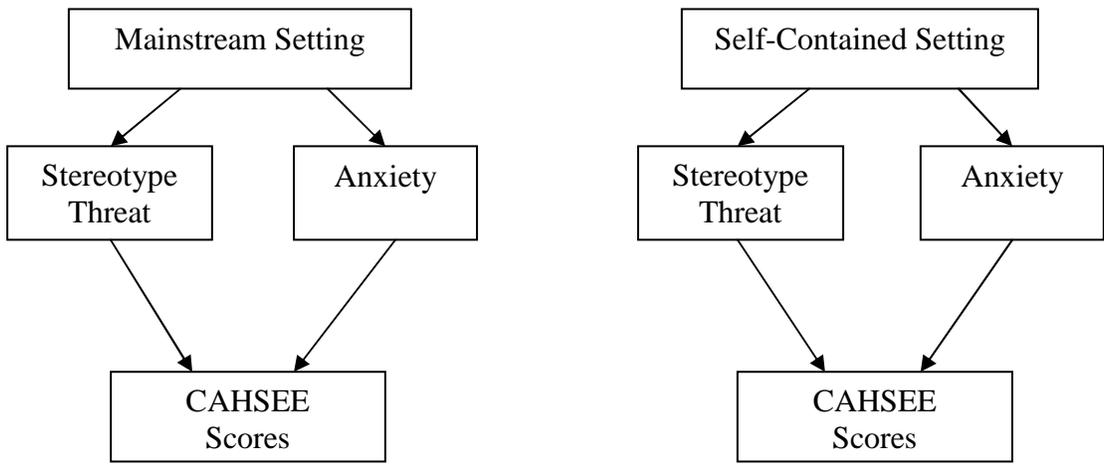


Figure 1. The impact of setting, stereotype threat, and anxiety on CAHSEE scores.

**Setting**

The study was conducted at three different high schools in three suburban school districts in Northern California. Demographic data on the schools were obtained for each of the school sites. Table 1 below provides a summary data of the schools' demographics.

Table 1

*Summary Demographics of Participating Schools*

	School Alpha	School Beta	School Gamma
Academic Performance Index	9	9	7
School Size	2897	1474	4150
Number of English Learners	88	59	595
Total Number of Students with Disabilities	216	125	340
Number of Socioeconomically Disadvantaged	559	224	1571
Number of Student Enrollment by Ethnicity:			
White	1087	754	430
African-American	161	129	413
Hispanic/Latino(a)	588	209	1257
Asian	665	254	962
Pacific Islander	16	18	187
Filipino	70	87	810
Alaska/Native American	5	10	11
Multi-Ethnic (Other)/Not Reported	305	23	80

Source: California Department of Education for fiscal year 2009-2010

The demographic data showed all three schools serve a similar percentage of students with disabilities, and that school Gamma had the highest student enrollment and serves more ethnically diverse students than the other two school sites. The three schools who participated in the study have Academic Performance Index (API) scores that range from 7 to 9. This is above the average API scores for California. School Alpha had more

multiple-ethnic or students who did not report their ethnicity (11%) compared to schools Beta and Gamma (2%), and less than 5% of the student population in schools Alpha and Beta were English learners.

### **Sample**

All participants in the study were recruited from three comprehensive, 9-12 public high schools located in a suburban community in Northern California's Bay Area. The participants utilized in the study were a convenience sample of students with disabilities enrolled in their sophomore year of high school, identified as having an Individualized Education Plan (IEP) by the school, and who took the CAHSEE for the first time in the spring administration of the test. There were no criteria for students' gender or ethnicity. The total number of potential participants from the three school sites was 135.

All 10<sup>th</sup> grade students with disabilities who took the CAHSEE for the first time in the spring administration were eligible to participate and were recruited to do so via an introduction of the study by the schools' special education staff and the researcher. The most common reason for excluding a potential participant was non-return of the consent forms. Participation in this study was voluntary and respondents were not anonymous to the researcher. Of the 135 targeted participants, a total of 66 students (approximately 49%) returned the consent forms and participated in the final study. The majority of the participants were from school Alpha ( $N=38$ ). School Gamma had the next largest number of participants ( $N=20$ ), whereas school Beta had the smallest number of participants ( $N=8$ ). Two students from school Beta were excluded from the study due to lack of information regarding their CST scores which brought the total number of participants to 64. Table 2 provides demographic data of the participants in the study.

Table 2

*Frequency Table of Demographics for Participants*

	<b>Mainstream</b> ( <i>n</i> )	<b>Self- Contained</b> ( <i>n</i> )
Total Participants	33	31
Gender		
Male	18	24
Female	15	7
Ethnicity		
White	16	9
African-American	5	5
Hispanic/Latino(a)	7	10
Asian	2	4
Pacific Islander/Filipino	2	0
Alaska/Native American	0	0
Multi-Ethnic (Other)	1	3
Primary Disability		
Specific Learning Disability	22	13
Emotional Disorder	0	0
Speech and Language	4	7
Visual Impairment	0	1
Hearing Impairment	1	0
Autism/Autistic-Like	4	2
Other Health Impaired	2	8
English Learner	5	8
Year Algebra 1 Taken		
8 <sup>th</sup> grade	2	0
9 <sup>th</sup> grade	25	13
10 <sup>th</sup> grade	5	18
Other	1	0
CAHSEE Pass Status	26	8

The participants were divided into two groups for data analysis, those who received math instruction in a mainstream (general education) setting and those who received math instruction in a self-contained (special education) setting. Within a mainstream setting, students with disabilities attend math class with students who do not

have disabilities and the instructor will have a credential in the subject being taught. In a self-contained class, students with disabilities attend math class with students who have disabilities and the instructor will have a credential in special education.

A total of 52% of the participants were in a mainstream setting and 48% were in a self-contained setting. Of the total participants none were classified as Alaska/Native American or had an emotional disorder. The majority of the participants (55%) in both settings were classified as having a specific learning disability. Students in the mainstream setting had an even amount of males and females, with almost half of the students classified as White/Caucasian. The majority of the participants in the self-contained setting was male (77%) and approximately one-third were classified as White/Caucasian. Approximately one-quarter of the students in the self-contained setting were EL ( $N=8$ ) compared to approximately 15% for the mainstream setting. Most students (76%) in the mainstream setting took Algebra 1 in their freshman year of high school. For the self-contained setting, less than half (41%) of the students took Algebra 1 as a freshman in high school with the majority (58%) of students taking the class in their sophomore year. The majority of the students in the mainstream setting passed the math section of the CAHSEE (79%) on the first administration; less than half (26%) in the self-contained setting passed. Students' SES was not obtained for this study as the data were not available.

### **Protection of Human Subjects**

The study received approval from the Institutional Review Board for the Protection of Human Subjects (IRBPHS) committee at the University of San Francisco (see Appendix A). The protection of the general welfare of human subjects consisted of

informing all potential participants and their guardians of the general purpose of the study. All the potential participants and their guardians were informed in writing that his or her child's participation was voluntary and there were no direct benefit for participating in the study. In addition, the participants and their guardians were informed that they had a right to withdraw from the study at any time and the participant's identity would be protected. Neither the content of the questionnaire nor the data collection methods were believed to impose any physical, psychological or other unknown harm to any participant in this study.

All data collected were kept in a secure place under lock and key. All electronic data were kept on a password protected file on an internal and external computer drive. The anticipated benefit of this study was the understanding of the relationship of course setting on stereotype threat, anxiety, and the score on the mathematics section of the CAHSEE for students with disabilities. Students were encouraged to participate in the study with the use of a researcher-funded raffle of a \$25.00 Visa gift card per school site. Students did not have to complete the questionnaire in order to be eligible for the raffle. One hundred percent of student participants were eligible to participate in the raffle including those who participated in the pilot study ( $N=96$ ). To be eligible for the raffle, students were asked to voluntarily provide their names on the coversheet of the questionnaire. One questionnaire was randomly selected from each school site in the final study to win the \$25.00 Visa gift card.

### **Instrumentation**

Three instruments were used in this study. The first instrument used was the mathematics section of the California High School Exit Exam (CAHSEE), followed by a

questionnaire to measure anxiety, and a questionnaire to indirectly measure stereotype threat. Each instrument is described below.

**Math section of CAHSEE.** As stated by the California Department of Education, the primary purpose of the CAHSEE is to significantly improve student achievement in public high schools and to ensure that students who graduate from public high schools in California can demonstrate grade level competency in reading, writing, and mathematics (California Department of Education, 2008a). The CAHSEE is a two part exam consisting of English/language arts and mathematics. The California Department of Education selected the nonprofit company, Educational Testing Services (ETS) to develop, administer, and score the CAHSEE.

The mathematics section of the CAHSEE addressed state mathematics content standards in grades six and seven and the first part of Algebra. The exam is composed of 92 multiple-choice questions (twelve of which are field test items and are not scored). The exam includes statistics, data analysis and probability, number sense, measurement and geometry, algebra and functions, mathematical reasoning, and Algebra 1. The number correct or raw scores on the CAHSEE are converted to a scale score in order to maintain comparability of scores across multiple test forms. The raw score to scale score conversion reflects the relationship between difficulty of individual test questions in each test form and the constant measure of achievement indicated by the reported scale scores (California Department of Education, 2008c). The procedure of converting the raw scores to scale scores involves scaling and equating. A scale score of 350 or higher is required to pass the CAHSEE mathematics section.

For students with disabilities, a local waiver of the CAHSEE requirement may be available if the student who took the CAHSEE using modifications received the

equivalent of a passing score. The modification must be written in the student's IEP for use on the CAHSEE, standardized testing, or for use during classroom instruction and assessments (California Department of Education, 2008a). However, for the purpose of this study, only the students' scores who took the test without the use of modifications were analyzed.

**Anxiety instrument.** The second measurement instrument was a student questionnaire to measure test anxiety with a cover letter provided (see Appendix B and C). The questionnaire was utilized because large amounts of data could be collected in a relatively short period of time and it is cost effective. The questionnaire is a self-reported measure which consists of seventeen 4-point Likert-scale items. The items were constructed and adapted from the State-Trait Anxiety Inventory, an existing test anxiety measure which has a reliability of .86 (Spielberger, Gorsuch, & Lushene, 1970). The language was changed to past tense as the instrument was given after the CAHSEE had been administered. The scale ranged from 1 (strongly disagree) to 4 (strongly agree). Several items on the questionnaire (#1, 2, 5, 7, 9, 10, 14, and 17) were reversed coded so that the scaled items are oriented in the same direction. The instrument was scored using total points. Each item was scored from 1 to 4 with total points ranging from 17 to 68. A low total number would indicate a lower level of anxiety whereas a high total number would indicate a higher level of anxiety. In the final study, Cronbach's Alpha for the seventeen anxiety questionnaire items was .917.

**Stereotype threat instrument.** The third measurement instrument was a student questionnaire (see Appendix D) which was used to indirectly measure stereotype threat. Stereotype threat had been measured using a variety of instruments, mainly surveys

produced by researchers (Hollis-Sawyer & Sawyer, 2008; Kellow & Jones, 2008; Ployhart, Ziegert, & McFarland, 2003) that measured anxiety, self-efficacy, as well as evaluation apprehension. This study indirectly measured stereotype threat using a researcher created self-reported instrument aimed at specifically addressing students with disabilities.

The questionnaire was adapted from measures of evaluation apprehension and self-efficacy, which can be mechanisms by which stereotype threat occurs (Spencer, Steele, & Quinn, 1999; Steele & Aronson, 1995). A 4-point Likert scale was used in the development of the 9 questionnaire items. The scale ranged from 1 (strongly disagree) to 4 (strongly agree). The instrument was scored using total points. Each item was scored from 1 to 4 with total points ranging from 9 to 36. The 4-point Likert scale method was selected to encourage students to choose to agree or disagree in an attempt to lessen their indecision on each item. A low total number would indicate a low level of stereotype threat and a high total number would indicate a high level of stereotype threat. In the final study, Cronbach's Alpha for the nine stereotype threat questionnaire items was .854.

### **Pilot Study**

A pilot test using thirty students with disabilities at a public high school in the Bay Area was conducted to test the student questionnaire which measured anxiety and indirectly measured stereotype threat. The pilot study was conducted in March, 2010 at a suburban public high school not included in the final study due to the earlier administration of the CAHSEE at the school site.

The researcher obtained permission from the school site administrator and the district to conduct the pilot study. The administrator referred the researcher to the special education department chair to provide a list of eligible students with disabilities, their

class periods and times, and a classroom to conduct the gathering of data. A brief overview about the study and its importance was provided to the pilot study participants by the researcher and special education department chair. Next, the anxiety and stereotype threat questionnaire were given to the students. The questionnaires took approximately 10 minutes to complete. Once done, the questionnaires were collected by the special education department chair and given to the researcher. Cronbach's alpha was used to estimate the reliability of the anxiety and the stereotype threat questionnaires.

The reliability for the 17 item anxiety questionnaire in the pilot was .925; thus, no changes were made. The original stereotype threat questionnaire consisted of 15 items. The reliability for the stereotype threat questionnaire in the pilot was .765. After the removal of six items through the process of item analysis, the reliability was increased to .827. Therefore, no additional modifications were made to the nine items following the pilot. The combined nine questionnaire items were thought to represent an indirect measurement of participants' level of stereotype threat at the time the study was conducted.

### **Validity Panel**

Prior to initiating the final study, a panel of experts in the field of education with an emphasis in special education was recruited to review the stereotype threat instrument to be used. Nine panel members were initially recruited to conduct the validity study; however, two never responded. The panel of experts consisted of program specialists, teachers, transition specialists, and an administrator; all who work directly or indirectly with secondary students with disabilities. Specific demographic information on members of the validity panel was obtained and recorded (see Appendix E AND F).

The validity panel packet was hand delivered to each member. The packet outlined the purpose of the questionnaire, and a brief definition of the theoretical influence for the study and the variables measured on the questionnaire. A set of prompts to guide the panel through the review of the questionnaire were provided in the packet (see Appendix G). The panel responded in writing to the set of prompts within the packet.

The seven panel members used their professional judgment and provided feedback on the face, internal, and content validity of the stereotype threat questionnaire protocol. Panel members were asked to write any comments or question they had regarding the face, internal, or content validity on the questionnaire items. Additionally, the panel members were asked to comment on the questionnaire's design, anticipated length of time of administration, appropriateness to and for the population of students targeted, and ease of administration. Each validity panel member completed the validity packet individually.

All the members of the validity panel completed the validity panel packet sent to them by the researcher. After the receipt of the completed validity panel packet, the validity panel participants received a \$10.00 Starbucks Gift Card with a personalized note card from the researcher thanking them for their feedback and support of this study.

Upon receipt of each reviewed packet, the comments were data processed into a template for ease of review (see Appendix H). Recommendations and comments made by each validity panel member were reviewed. The following narrative represents the modifications made to the original questionnaire based on the validity panel members' comments and recommendations.

The opening statement "Because I have a disability" was removed as a beginning prompt. To increase the likelihood that students would be able to comprehend and

respond to each question and for ease of reading, the stem “because I have a disability” was placed at the beginning or ending of a statement wherever applicable. Question 15 on the first page of the questionnaire was eliminated as it was a duplicate of question 7. More space was added between questions 7, 8, 13, and 14 on the second page of the questionnaire for consistency. Questions 3 and 15 on the second page of the instrument were reworded for clarity. Additionally, a period was added after each statement on the second page of the questionnaire.

As a result of the validity panel review and the pilot study reliability results, the formatting of the stereotype threat questionnaire was changed to increase the readability for high school students with learning disabilities, and for presentation purposes. Upon completion of all revisions, the list of questionnaire items measuring anxiety and stereotype threat previously presented in this chapter was reviewed and revised to reflect the new format of the questionnaire.

## **Procedures**

The researcher applied and was granted permission to conduct research from the University of San Francisco Institutional Review Board for the Protection of Human Subjects (IRBPHS) and the public school districts’ research offices. Letters, emails, and/or telephone calls to school administrators of several high schools in the Bay Area whose school API scores were between 7-10 were initiated by the researcher to briefly describe the study and ask for their support in the form of consent to conduct the study. Due to time constraints, the first three letters of support for the study received from the school or district administrators were the only schools selected for participation in the study. The final sample from which the participants were drawn was considered to be that of convenience.

**Data collection.** The researcher went into the schools to meet the special education teachers, discuss the purpose of and need for the study, and to discuss the administration process and procedures for securing and tracking parental permission forms. Permission from the school site administrator or district administrator was obtained in writing prior to selecting the samples and initiating any part of the study (see Appendix I). Each administrator referred the researcher to the special education department chair to provide a list of eligible students with disabilities, their class periods and times, and a classroom to conduct the gathering of data.

With the department chair's verbal permission, the researcher briefly introduced the study and its purpose to the eligible students in each classroom. The researcher distributed the parental permission documents to the students that were present and informed them that the researcher would be returning to their site to pick up the parental permission and administer the questionnaire after they have taken the CAHSEE. If the student was not present the researcher gave the parental permission documents to the department chair to distribute, to collect each of the returned permission slips, and to secure them in a manila envelope for pick up by the researcher.

Two copies of the Parental Agreement and Consent forms (see Appendix J) and a copy of the Research Subject's Bill of Rights (see Appendix K) was made for each student eligible from each site. One copy was given to the student and one copy was mailed home. Students who returned, or mailed/faxed the completed, approved, and signed consent form from each school site were eligible to participate in the study. The department chair's informed the researcher as to when each school's administration of the CAHSEE occurred. A second set of parental consent forms were mailed or taken to the department chairs if needed for any student who had not returned a parental consent form

signed by his or her parent(s) or guardian(s). The researcher's contact information was printed on the consent form in case parents had questions prior to signing the consent form. Four students received \$25.00 Visa gift cards in four independent raffles held for participants (one from each school site and one from the pilot study). Visa gift card winners were randomly selected from all participants at each school site and the winnings were hand delivered to the students at their respective schools.

**Questionnaire administration.** After the administration of the mathematic section of the CAHSEE by the school administration, within a three week period, students with disabilities were asked to complete the anxiety and stereotype threat questionnaires. The day of the questionnaire administration, the researcher reported to the school office with the number of questionnaires for each consent form received at each school. Students who had a signed consent form were taken to a quiet room on campus in the special education department where the researcher was located to complete the questionnaire during their academic or advising period in special education. Introduction of the researcher was initiated with all participants.

Participants were briefed on the purpose of the study and thanked in advance for their time. The participants were given another opportunity to ask questions related to the study. The researcher handed out the questionnaire with the assistance of the department chair when available. All questionnaire directions and all questionnaire items were read by the participant. The directions or any word was read aloud to the participant only if requested. The researcher quietly answered any questions the participants had during the questionnaire's administration so the other participants could continue without interruption. The same procedures were followed for any participant that was absent on the initial administration of the questionnaire. The administration process took

approximately 10 – 15 minutes. The administration of the questionnaires did not interfere with the regular school functioning nor the students' normal routines. The classrooms used in the study were in special education settings. Students in mainstream settings were administered the questionnaire during their academic support classes with their special education teacher to minimize the disruption to the students' daily routines.

After the administration of the questionnaire, the questionnaires were collected by the researcher or special education department chair. The participants were thanked a second time and given a short opportunity to ask the researcher questions related to the student. Once all questions were answered, if any, the participants were told that the raffle would be held within the next few weeks and that the winner would be notified. The participants and department chairs were thanked a final time by the researcher and the administration of the questionnaire was concluded. To ensure that confidentiality was maintained, the questionnaires were taken to the researcher's home office and kept in a locked desk drawer maintained by the researcher until needed for data entry and analysis. Once the data were analyzed, the questionnaires were returned to the locked desk drawer in the researcher's office.

**School record review.** Once the results from the CAHSEE were received by the school district from the California state department of education, the researcher was given access to the school computerized data information system or the students school cumulative folder to obtain data related to the background variables (e.g. prior math ability) and demographic information which was used to complete the student data sheet (see Appendix L). Since the participants records were confidential and could not be taken off school grounds, the researcher completed the student data sheets on the school or

district site. The data from the questionnaire, code sheet, and the scores on the mathematics section of the CAHSEE were coded and quantified.

The collection of data from the anxiety and stereotype threat questionnaires began in March of 2010 after the first spring administration of the CAHSEE. The data collection regarding the scores on the mathematics section of the CAHSEE began once the school districts received notification of the test results in June, which occurred approximately seven to eight weeks after the administration of the CAHSEE. The data collected were kept secure in the researcher's home office under lock and key. All electronic data were kept on an internal and external computer drive under a password protected file. The data gleaned from the review of student records were entered into the same SPSS file containing the participants' individual questionnaire responses. The data were summarized and the findings are discussed in Chapter IV of this study.

**Data analysis.** Descriptive and inferential statistics were applied to the questionnaire data collected from all the students to determine whether statistically significant means differences existed between students with disabilities in mainstream and self-contained settings on anxiety, stereotype threat, and the math section of the CAHSEE. The Statistical Package for the Social Sciences (SPSS) software versions 17 and 18 with alpha set at .05, unless otherwise noted, were utilized to analyze the data. Descriptive statistics (frequencies, percentages, means, and standard deviations) was used to present quantitative data gathered from the school database system, questionnaire, and coding sheet.

To answer the three research questions, a one-way analysis-of-covariance (ANCOVA) was performed to determine statistically significant mean differences in test scores between the student groups after controlling for differences on the covariate. The

researcher used the principal components analysis (PCA) to transform the set of correlated variables (CST mathematics scores for grades 8<sup>th</sup> and 9<sup>th</sup> and Algebra 1 math course grades) into a single covariate described as prior math ability. The assumptions that underlie a one-way ANCOVA are normal distribution of the population, equal population variances, independence of populations, assumption of linearity, and homogeneity-of-regression.

The assumptions of normal distribution and equal variance of populations were met by utilizing a large sample size  $N = 33$  and  $N = 31$  for mainstream and self-contained groups respectively. Equality of variance was also tested using the Levene's Test. The results indicated no statistical significance. The homogeneity of regression slopes was tested using SPSS. The results are discussed after each question. The participants are separated based on the setting by which they receive math instruction without interaction. Thus, the assumption of independence of populations was met. The assumption of linearity was met by conducting a visual inspection of the scatterplots.

### **Summary**

This causal comparative study explored the educational setting, test anxiety, and stereotype threat which may contribute to the scores on the California High School Exit Exam mathematics section for students with disabilities who take the test for the first time in their sophomore year of high school. Causal comparative designs are frequently used in research which utilizes intact groups, no random assignment of participants to a group, allowing for some generalizations to be made about a population. Specific procedures for the development of the questionnaire were presented in this chapter as well as specific methods used in conducting this study.

The stereotype threat questionnaire used in this study was reviewed by a panel of experts and pilot tested in a suburban public high school. The pilot study suggested that the items used to indirectly measure stereotype threat had a relatively high internal consistency. All quantitative analyses were conducted using SPSS Versions 17 and 18.

## CHAPTER FOUR

### RESULTS

This chapter presents the results of data analyses conducted to answer the three research questions presented in this study, and is organized into five sections. First, an introduction to the chapter is provided followed by descriptive statistics. Next, each of the three research questions are presented and addressed in numerical order. Lastly, the chapter concludes with a summary of the results.

This study examined whether significant mean differences existed between students with disabilities in mainstream and self-contained settings on anxiety, stereotype threat, and the scores on the math section of the CAHSEE. Prior to all analyses, descriptive statistics were obtained using SPSS Version 17 to investigate the means, standard deviations, normality, linearity and homoscedasticity for each of the variables in the study for the mainstream ( $n = 33$ ) and self-contained ( $n = 31$ ) samples.

#### **Descriptive Statistics**

Table 3 presents the means and standard deviations of CAHSEE math scores, STT, anxiety, 9<sup>th</sup> grade point average for fall and spring, 8<sup>th</sup> and 9<sup>th</sup> grade CST math score, Algebra 1 grade for fall and spring semester, and PMA by setting. All tests of statistical significance were conducted at the .05 level of significance.

Table 3

*Means, Standard Deviations (SD), and Sample Sizes (n) of California High School Exit Exam (CAHSEE) Math Scores, Stereotype Threat (STT), Anxiety, Overall 9th Grade Point Average (GPA), and Prior Path Ability (PMA) Components Scores for Mainstream and Self-Contained Groups*

	<u>Mainstream</u> <i>n</i> = 33		<u>Self-Contained</u> <i>n</i> = 31	
	Mean	SD	Mean	SD
CAHSEE	364.48	27.61	330.97	18.43
STT	2.16	.65	2.37	.47
Anxiety	2.34	.55	2.44	.53
Overall GPA Fall	2.80	.73	2.35	.98
Overall GPA Spr.	2.60	.88	2.33	.98
CST 8	314.10	54.22	265.81	44.27
CST 9	307.70	42.82	262.19	42.42
Algebra Grade Fall	2.24	1.17	1.81	1.38
Algebra Grade Spr.	2.15	1.35	2.06	1.15
PMA	.36	1.00	-.39	.86

The mean scores for student in mainstream settings on CAHSEE math scores were higher and more varied than those of the self-contained settings. Students in the mainstream setting ( $M = 364.48$ ,  $SD = 27.61$ ) on average had higher mean scores on the CAHSEE math section than those in the self-contained setting ( $M = 330.97$ ,  $SD = 18.43$ ). Results on stereotype threat indicate that on average, students who received math instruction in a self-contained setting had a higher mean score ( $M = 2.37$ ,  $SD = .47$ ) than the students who received math instruction in a mainstream setting ( $M = 2.16$ ,  $SD = .65$ ). The results suggest that students in self-contained settings experience more stereotype threat than those in mainstream settings. The results on the anxiety measure indicate that on average, students who receive math instruction in a self-contained setting have a higher mean score ( $M = 2.44$ ,  $SD = .55$ ) than students who receive math instruction in a

mainstream setting ( $M = 2.34$ ,  $SD = .53$ ). The results suggest that students in a mainstream setting experience more anxiety than students in a self-contained setting.

Overall GPA for fall and spring semesters of the participants' freshman year suggest that students in mainstream settings had higher fall GPA ( $M = 2.80$ ,  $SD = .73$ ) and spring GPA ( $M = 2.60$ ,  $SD = .88$ ) than students in self-contained settings for fall ( $M = 2.35$ ,  $SD = .98$ ) and spring ( $M = 2.33$ ,  $SD = .98$ ). While there appears to be a drop in GPA from the fall to spring semesters for students in the mainstream setting, the students' in the self-contained setting GPA for fall and spring remained relatively the same.

Results of the mean differences for the components of the PMA scores suggest that on average, students in the mainstream setting had higher CST mean scores for both 8<sup>th</sup> and 9<sup>th</sup> grades than the students in self-contained settings. The Algebra mean grade scores for fall semester was higher for the mainstream setting than for the self-contained setting, but the Algebra mean grade scores for spring semester was similar for both settings.

Correlation coefficients were computed among the 10 variables in the study. Using the Bonferroni approach to control for Type I error across the 45 correlations, a  $p$  value of less than .005 was required for significance. The results presented in Table 4 show that 24 out of 45 correlations were statistically significant. The correlation of CST 9 and STT was significant,  $r(62) = -.16$ ,  $p < .005$  and the correlation of CAHSEE and anxiety was significant,  $r(62) = -.37$ ,  $p < .005$ . In general, the results suggest that CAHSEE scores are positively correlated with PMA, negatively correlated with anxiety, and negatively correlated with STT although not significantly.

Table 4

*Correlations among Anxiety, Stereotype Threat (STT), Prior Math Ability (PMA) Components Scores, Overall 9th Grade Point Average (GPA), and California High School Exit Exam (CAHSEE) Math Scores (N = 64)*

	Anxiety	STT	PMA	CST 8	CST 9	Algebra Grade Fall	Algebra Grade Spring	GPA Fall	GPA Spr
STT	.60*								
PMA	-.14	-.18							
CST 8	-.27	-.22	.77*						
CST 9	-.13	-.16*	.81*	.75*					
Algebra Grade Fall	-.05	-.09	.79*	.35*	.40*				
Algebra Grade Spr	.04	-.06	.69*	.22	.29	.67*			
GPA Fall	-.07	.04	.67*	.34	.45*	.68*	.59*		
GPA Spr	-.07	.09	.65*	.34	.44*	.60*	.64*	.85*	
CAHSEE	-.37*	-.29	.66*	.73*	.75*	.32	.19	.42*	.34

\*  $p < .005$

A principal components analysis (PCA) was used to transform the set of correlated variables (CST mathematics scores for grades 8<sup>th</sup> and 9<sup>th</sup> and Algebra 1 math course grades) into a single variable described as prior math ability. A component analysis using SPSS was conducted on the four scores to transform the variables into one variable called prior math ability. Prior math ability was used as the covariate.

The loadings for CST8, CST 9, Algebra Grade Fall and Algebra Grade Spring on the first principal component were .77, .81, .79, and .69, respectively. The results of the PCA indicates that the corresponding factors used to derive the PMA variable are above

(greater than) the mean for the students in the mainstream setting and below (less than) the mean for the students in the self-contained setting.

A multiple regression analysis was conducted to evaluate how well the independent variables and covariate measures (anxiety, stereotype threat, and prior math ability) predicted CAHSEE math scores for students with disabilities. The linear combination of all predictors was significantly related to the CAHSEE math scores,  $R^2 = .51$ ,  $F(3, 60) = 20.99$ ,  $p < .01$ . The multiple correlation coefficient was  $R = .71$  and  $R^2 = .51$ , indicating that approximately 51% of the variance of the CAHSEE math scores in the sample can be accounted for by the linear combination of the independent variables and covariate measures (see Table 5 below). Two of the beta weights were negative and two of the three were statistically significant ( $p < .05$ ). The results indicated that if anxiety was increased then CAHSEE math scores would decrease. In addition, if there was an increase in prior math ability then CAHSEE math scores would also increase.

Table 5

*Summary of Multiple Linear Regression of California High School Exit Exam Scores onto Anxiety, Stereotype Threat (STT) and Prior Math Ability (PMA)*

Variables	B	SE B	t	p
Anxiety	-.27	6.06	-2.39	.02
STT	-.02	5.68	-.16	.88
PMA	.62	2.65	6.74	.00

\* $p < .05$

### **Analysis Related to Research Question 1**

The first research question in the study asked whether there was a significant difference in the scores on the mathematics section of the CAHSEE for 10<sup>th</sup> grade students with disabilities enrolled in mainstream or self-contained settings for math instruction when prior math ability (as measured by Algebra 1 grade, 8<sup>th</sup> and 9<sup>th</sup> grade CST scores on Math) was controlled. The researcher used the principal components analysis (PCA) to transform the set of correlated variables (CST mathematics scores for grades 8<sup>th</sup> and 9<sup>th</sup> and Algebra 1 math course grades) into one variable described as prior math ability. Prior math ability (PMA) was used as the covariate.

A one-way analysis of covariance (ANCOVA) was conducted to address the first question. The independent variable, setting, included two levels: mainstream (group 1) and self-contained (group 2). The dependent variable was the students' math score on the CAHSEE and the covariate was the students' PMA scores. A preliminary analysis evaluating the homogeneity-of-regression (slopes) assumption indicated that the relationship between the covariate and the dependent variable did not differ significantly as a function of the independent variable,  $F(1,60) = 3.82, p = .06$ , partial eta squared = .06. Therefore, the assumption was met. The ANCOVA was statistically significant,  $F(1, 63) = 18.46, p < .01$  (See Table 6). The strength of relationship between the setting factor and dependent variable was medium, as assessed by a partial eta squared, with the setting accounting for 23% of the variance of the dependent variable, holding constant the students' prior math ability.

Table 6

*Analysis of Covariance Summary for California High School Exit Exam by Setting*

Source	Sum of Squares	<i>df</i>	Mean Square	<i>F</i>	<i>P</i>
PMA	11755.67	1	11755.67	31.43	.000
Setting	6903.15	1	6903.15	18.46	.000
Error	22815.54	61	374.03		
Corrected Total	52528.00	63			

\*  $p < .05$ **Analysis Related to Research Question 2**

The second research question asked whether a statistically significant mean difference existed in stereotype threat between 10<sup>th</sup> grade students with disabilities enrolled in mainstream or self-contained settings for math instruction when prior math ability (as measured by Algebra 1 grade, 8<sup>th</sup> and 9<sup>th</sup> grade CST scores on Math) has been controlled.

A one-way analysis of covariance was conducted to answer the second question. The independent variable, setting, involved two levels: mainstream and self-contained. The dependent variable was the measure of stereotype threat scores and the covariate was PMA. A preliminary analysis evaluating the homogeneity-of-regression (slopes) assumption indicated that the relationship between the covariate and the dependent variable did not differ significantly as a function of the independent variable,  $F(1, 60) = .78, p = .38$ , partial eta squared = .01. Therefore, the assumption was met. The results of the ANCOVA was not significant,  $F(1, 61) = .92, p = .34$  (see Table 7).

Table 7

*Analysis of Covariance Summary for Stereotype Threat by Setting*

Source	Sum of Squares	<i>df</i>	Mean Square	<i>F</i>	<i>p</i>
PMA	.29	1	.29	.88	.35
Setting	.30	1	.30	.92	.34
Error	20.14	61	.33		
Corrected Total	21.09	63			

\* $p < .05$

### **Analysis Related to Research Question 3**

The third research question in the study asked whether a statistically significant mean difference existed in anxiety for 10<sup>th</sup> grade students with disabilities enrolled in mainstream or self-contained settings for math instruction when prior math ability (as measured by Algebra 1 grade, 8<sup>th</sup> and 9<sup>th</sup> grade CST scores on Math) has been controlled.

A one-way analysis of covariance was conducted to answer the third question. The independent variable, setting, involved two levels: mainstream and self-contained. The dependent variable was the measure of anxiety scores and the covariate was PMA. A preliminary analysis evaluating the homogeneity-of-regression (slopes) assumption indicated that the relationship between the covariate and the dependent variable did not differ significantly as a function of the independent variable,  $F(1, 60) = .70, p = .13$ , partial eta squared = .04. Therefore, the assumption was met. The ANCOVA was not significant,  $F(1, 61) = .10, p = .76$  (see Table 8).

Table 8

*Analysis of Covariance Summary for Anxiety by Setting*

Source	Sum of Squares	<i>df</i>	Mean Square	<i>F</i>	<i>p</i>
PMA	.24	1	.24	.82	.37
Setting	.03	1	.03	.10	.76
Error	17.94	61	.29		
Corrected Total	18.33	63			

\* $p < .05$

**Summary**

This study explored the CAHSEE math scores, anxiety, and stereotype threat of suburban secondary students with disabilities in mainstream and self-contained setting for math instruction. The investigation into outcomes on the CAHSEE math scores studied the differences in anxiety, and stereotype threat experienced between both groups. Prior math ability was a controlled factor in relation to this study.

An analysis of covariance was conducted to answer the three research questions. Table 9 shows a summary of the means and adjusted means for each variable by setting. Students with disabilities in the mainstream setting had higher adjusted mean scores for CAHSEE math and lower adjusted mean scores for anxiety and stereotype threat.

Table 9

*Summary of Analysis of Covariance Mean and Adjusted Mean for California High School Exit Exam (CAHSEE), Anxiety, and Stereotype Threat (STT) by Setting*

Variable	Setting	n	Mean	Adjusted Mean
CAHSEE	Mainstream	33	364.48	359.12
	Self-Contained	31	330.97	336.68
Anxiety	Mainstream	33	2.34	2.37
	Self-Contained	31	2.44	2.41
STT	Mainstream	33	2.16	2.19
	Self-Contained	31	2.37	2.34

\* $p < .05$

The main findings were as follows: First, students in mainstream setting had higher math scores on the CAHSEE than the students taught mathematics in self-contained settings. Second, students in a mainstream and self-contained setting indicated little stereotype threat although scores indicate a higher level for students in the self-contained setting than students in the mainstream setting. Third, students in self-contained settings indicated some anxiety, but slightly less than students in a mainstream math setting.

A statistically significant difference was found for the adjusted mean CAHSEE math scores between the mainstream and self-contained settings after prior math ability was controlled. The strength of relationship between settings and the CAHSEE math scores, as assessed by partial eta squared, was medium (Cohen, 1992), with the setting accounting for 23% of the variance of the dependent variable, holding constant the prior math ability.

**CHAPTER FIVE**  
**SUMMARY, LIMITATIONS, DISCUSSION, CONCLUSIONS**  
**AND IMPLICATIONS**

This chapter is organized into six sections. The chapter begins with the purpose of the study, theoretical rationale, background and need, as well as the significance of the study. Second, the chapter includes a summary of findings and a discussion of the study's limitations. Next is a discussion of the findings as they relate to each research question, conclusions, and lastly are implications for future research and practice.

This causal comparative study examined the impact of setting, stereotype threat, and anxiety on the math scores of the California High School Exit Exam for students with disabilities who receive math instruction in a mainstream or self-contained setting. This study examined the hypothesis that mainstream settings help reduce anxiety and stereotype threat, thereby increasing performance in a high stakes testing situation.

The theoretical rationale underlying this study is stereotype threat theory (Steele & Aronson, 1995). Stereotype threat is the belief or threat of confirming a negative stereotype about a group to which one belongs (Steele, 1997; Steele & Aronson, 1995). Stereotype threat can affect anyone with a group identity to which a negative stereotype exists. Stereotype threat, which can be signaled by the mere recognition or association with a negatively stereotyped group, may be a factor in a student's low academic performance (Steele, 1997). Thus, stereotype threat has been seen as a factor in explaining achievement gaps in standardized testing and applies to any group membership where underperformance could confirm a stereotypical expectation.

The study was conducted in three suburban high schools from three different districts within the Bay Area of Northern California. A sample of 64 students with

disabilities and identified as having an Individualized Education Plan (IEP) who took the CAHSEE for the first time in the spring administration of the test was utilized. There were no criteria for students' gender or ethnicity. Slightly over half, 52%, of the participants received their math instruction in the mainstream setting while 48% received their math instruction in a separate self-contained setting.

Three instruments were used in this study: results of the mathematics section of the CAHSEE, a student questionnaire to measure test anxiety, and a student questionnaire which was used to indirectly measure stereotype threat. After the administration of the mathematic section of the CAHSEE by the school administration, within a three week period, the participants were asked to complete the anxiety and stereotype threat questionnaires. The anxiety and stereotype threat questionnaires were self-reported measures. The items used to measure anxiety were constructed and adapted from an existing test anxiety measure (Spielberger, Gorsuch, & Lushene, 1970). The questionnaire used to indirectly measure stereotype threat was adapted from measures of evaluation apprehension and self-efficacy, which can be mechanisms by which stereotype threat occurs (Spencer, Steele, & Quinn, 1999; Steele & Aronson, 1995).

This study is important for two reasons. First, exit exams have a tremendous impact on the post high school achievement of students with disabilities because high school exit exams determine the likelihood that a student will receive a high school diploma in California. For students with disabilities, the possession of a high school diploma will enable them to enter the work force at a high wage and continue their education on the job or obtain additional education in a supported college program (Albrecht & Joles, 2003).

Second, research with regards to students' with disabilities academic achievement based on instructional setting has been mixed. A study conducted by Rea, McLaughlin, and Walther-Thomas (2002) found that students with disabilities achieve better outcomes when they have been included in a mainstream setting than in a self-contained setting for mathematics. In contrast, Fore III, Hagan-Burke, Burke, Boon, and Smith (2008) found little support for one setting over another. This study supports and enhances previous research related to setting, stereotype threat, and anxiety in relation to high stakes standardized test results for students with disabilities.

To address the two concerns, this study examined the differences in student performance on the California High School Exit Exam (CAHSEE) between 10<sup>th</sup> grade students with disabilities who were enrolled in a mainstream or self-contained setting for math instruction. Following the collection of data from the students, the data were analyzed in SPSS ®. Descriptive and inferential statistics were obtained. A correlation matrix and a multiple regression analysis were conducted along with three one-way analysis of covariance (ANCOVA).

### **Summary of Findings**

To answer the three research questions, a one-way analysis-of-covariance (ANCOVA), was performed to determine significant differences in adjusted mean scores between students in a mainstream and students in a self-contained setting for math instruction. The researcher used the principal components analysis (PCA) to transform the set of correlated variables (CST mathematics scores for grades 8<sup>th</sup> and 9<sup>th</sup> and Algebra 1 math course grades) into one variable described as prior math ability. Prior math ability was used as a covariate. A multiple regression analysis was done to explain relationships

among variables. Anxiety was found to be significantly correlated to CAHSEE math scores.

A statistically significant adjusted mean difference was found at the .05 level of significance on the CAHSEE math scores between the mainstream and self-contained settings after prior math ability was controlled. The mean scores for students with disabilities in mainstream settings on CAHSEE math scores were significantly higher on average and more varied than those of the students with disabilities in the self-contained settings. The strength of relationship between settings and the CAHSEE math scores, as assessed by partial eta squared, was medium (Cohen, 1992), with the setting accounting for 23% of the variance of the dependent variable, holding constant the prior math ability.

Although the mean difference was not statistically significant for stereotype threat or anxiety, the results of the ANCOVA for stereotype threat indicated that on average, students who received math instruction in a self-contained setting had a higher mean score than the students who received math instruction in a mainstream setting. The results on the anxiety measure indicated that on average, students who receive math instruction in a self-contained setting have a higher mean score than students who receive math instruction in a mainstream setting.

### **Limitations**

Several limitations exist relative to this study. One limitation to this study is the setting in which the study was conducted. A small number of students surveyed in this study ( $N = 64$ ) were from three suburban school districts in Northern California's Bay Area. Therefore, the ability to make broad generalizations and/or inferences is limited to the Bay Area's suburban areas and other communities like it.

Another limitation to this study was the procedure for administering the questionnaire. The questionnaire was a self-reported measure. Participants may have responded in a way that they believe would reflect them more positively. Therefore, participants may not have answered the questions truthfully. The time between the CAHSEE test administration and the administration of the questionnaires is also a limitation to this study. Since the questionnaires were administered within two weeks after the test, students may not have experienced stereotype threat or anxiety in the same manner as they would have if the questionnaires were given just after the CAHSEE was administered.

Additionally, the natural clinical setting for this study was not readily conducive to a true random assignment of students to the groups because it was necessary to work with the intact groups already formed by the schools. There is a possibility that the students who were in the self-contained group were less motivated than the students in the mainstream group. There is also a possibility that there were some unknown differences between the mainstream and self-contained group, which contributed to a self-selection bias, and in turn, may have interfered with the outcomes.

Finally, multiple-choice items rather than open-ended response items may have prompted participants to provide answers that they might not have generated independently and the limited number of items may not have adequately addressed the breadth of information on the topic of stereotype threat and anxiety as it relates to the CAHSEE for students with disabilities. The next section discusses the major findings of this study as they relate to stereotype threat, anxiety, and settings' impact on California High School Exit Exam scores for students with disabilities.

## **Discussion of Findings**

The primary purpose of this study was to explore the impact of setting, stereotype threat, and anxiety on the California High School Exit Exam (CAHSEE) math scores for first time examinees with disabilities. The goal of special education is to have the delivery of education in the best possible environment for learning (Kavale & Forness, 2000). This delivery typically occurs in a mainstream classroom setting or self-contained setting. Advantages of self-contained settings over mainstream settings include smaller class sizes and more opportunities for individualized instruction to meet the needs of the students with disabilities. Advantages of mainstream settings include higher self esteem, better social skills, and access to the general curriculum.

The first question addressed whether a statistically significant mean difference existed in the scores on the mathematics section of the CAHSEE for 10<sup>th</sup> grade students with disabilities enrolled in mainstream or self-contained settings for math instruction when controlling for prior math ability (as measured by Algebra 1 grade, 8<sup>th</sup> and 9<sup>th</sup> grade CST scores on Math). In this study, a statistically significant mean difference on the math section of the CAHSEE was found between students with disabilities who receive math instruction in a mainstream setting and students with disabilities who receive math instruction in a self-contained setting when prior math ability was controlled. This finding is consistent with results that suggest that students in a self-contained setting for academic instruction do not equate to improved scores or pass rates on standardized tests (Rea, McLaughlin, & Walther-Thomas, 2002; Signor-Buhl, LeBlanc, & McDougal, 2006).

Students with disabilities educated in mainstream settings did achieve higher mean standard scores on the math section of the CAHSEE than students with disabilities

educated in self-contained settings. The result is consistent with the suggestion that while the field of special education evolved to serve students with specific needs, data on self-contained settings revealed non-satisfactory in terms of school achievement or long-term benefits (Rea, McLaughlin, & Walther-Thomas, 2002). This study suggests that students with disabilities can achieve academic success in general education classrooms.

Several factors may exist or explain why students with disabilities in the mainstream setting for math instruction did better on the CAHSEE. There may be a higher level of teacher expectation for student achievement in the mainstream setting. Students with disabilities in a mainstream setting are expected to learn the same material as their non-disabled peers, keep pace with the classroom instruction and homework, and achieve at the same level. Students with disabilities are expected to show their understanding of academic content through the same types of assessments or activities as students without disabilities, but may be given accommodations or modifications as needed and stated on the student's individualized education plan.

Students with disabilities in the mainstream setting may have been more prepared for the CAHSEE math test than students in the self-contained setting. In the mainstream setting, students with disabilities have more access to the general curriculum as opposed to the self-contained setting. The curriculum used in the mainstream setting is based on grade-level topic of the class whereas the self-contained setting uses curriculum to address deficits in the student's learning. All or a combination of the above reasons may have accounted for the higher level of outcome results on the CAHSEE math section for students with disabilities in the mainstream setting as opposed to the self-contained setting.

This study's results are also consistent with results that suggest that students with disabilities who are educated in inclusive settings will achieve at the same rate or somewhat better than students with disabilities educated in self-contained setting. Signor-Buhl, LeBlanc, and McDougal (2006) found that after controlling for IQ, students in mainstream classrooms performed significantly better on individual measures of reading achievement, English language arts assessment, and similar to students who were in self-contained classes on individual measures of math achievement. Students in mainstream settings performed slightly better on individual measures of math achievement, but not to any significant degree.

Although mainstream settings allows for access to the general curriculum, the students with disabilities may have received less teacher time and attention need for them to achieve at a statistically significant higher level above the students with disabilities in the self-contained settings. More time and attention to academic tasks can be effective in increasing students' academic skills thereby increasing the students' abilities used in the general educational environment. If students with disabilities in the mainstream setting where given needed time and possibly more directed instruction necessary, their performance on the CAHSEE may have resulted in a statistically significant difference between the two groups.

Other research related to the study of optimal setting for educational benefit for students with disabilities are inconsistent with this study. Fore III, Hagan-Burke, Burke, Boon, and Smith (2008) examined differences in academic achievement based on classroom placement for students with specific learning disabilities in secondary classroom settings. Their study found no statistically significant differences in the academic performance of students with disabilities for math between settings. The

implication of their findings which suggests little support for one class placement over the other is in conflict with this study's findings.

Several reasons may exist for why there was not a statistically significant difference found between the settings. First, the test given in Fore III et al. (2008) was a timed test whereas the CAHSEE is not timed. Students with disabilities who require additional processing time to complete an academic task did not have that opportunity which may have contributed to the non-significant difference in test performance between both settings. Second, the test used in Fore III et al (2008) is intended for educators who instruct and assess students exhibiting academic difficulties. As such, the element of a high stakes test was not present. Again this study found that setting did have an impact on student performance and as such warrants further examination.

The second question addressed whether a statistically significant mean difference existed in stereotype threat for 10<sup>th</sup> grade students with disabilities enrolled in mainstream or self-contained settings for math instruction when controlling for prior math ability (as measured by Algebra 1 grade, 8<sup>th</sup> and 9<sup>th</sup> grade CST scores on math). This study was based on Steele and Aronson (1995) theoretical construct of stereotype threat theory which posits that achievement is adversely affected by the belief or threat of confirming a negative stereotype about a group to which one belongs. Stereotype threat, which can be signaled by the mere recognition or association with a negatively stereotyped group, may negatively impact a student's academic performance (Steele, 1997). While stereotype threat has been widely used in research with various groups, it had not been empirically studied with students with disabilities at the high school level. In this study, a statistically significant mean difference was not found for stereotype

threat between the 10<sup>th</sup> grade students with disabilities enrolled in mainstream and self-contained settings for math instruction.

Findings in this study are consistent with Cadinu, Maass, Lombardo, and Frigerio's 2006 study on stereotype threat and its effect on academic performance. In their study, findings revealed that one's group performance relative to his or her abilities will be depressed because he or she is fighting against the possibility of failing the test. This suggests that an individual's performance under a stereotype threat situation will perform worse than those not under stereotype threat.

This study found that students in self-contained settings had a higher level of stereotype threat and a lower level of performance than students with disabilities in a mainstream setting although the one-way analysis of covariance result did not conclude a statistically significant adjusted mean difference. Although the results were not statistically significant between the two groups, results on the CAHSEE for students in both the mainstream and self-contained settings follow the theory of stereotype threat. This implies that students in the mainstream setting have the same level of negative group identification concern as do students with disabilities in self-contained settings, but their test scores are not depressed as a result. This is in contrast with this study's hypothesis that because students with disabilities are in a mainstream setting with other students who do not have a disabilities that they would feel more threat of being negatively stereotype based on the results of a high stakes assessment like the CAHSEE.

This result may have occurred because students with disabilities have been in the mainstream setting all year; they may have felt more comfortable about their abilities and themselves since they are not separated from the general population for math instruction. Also being in a classroom with their non-disabled peers, the interaction may have

fostered an understanding and awareness of academic challenges faced by students with disabilities. Students with disabilities may have developed more confidence and a higher level of self-efficacy, feeling equal to their peers who do not have a disability. In self-contained settings the interaction is with other students with disabilities, therefore, there is little opportunity for understanding and awareness of differences.

The third question of this study addressed whether a statistically significant mean difference existed in anxiety for 10<sup>th</sup> grade students with disabilities enrolled in mainstream or self-contained settings for math instruction when controlling for prior math ability (as measured by Algebra 1 grade, 8<sup>th</sup> and 9<sup>th</sup> grade CST scores on math). In this study, a statistically significant mean difference was not found for anxiety between the 10<sup>th</sup> grade students with disabilities enrolled in mainstream and self-contained settings for math instruction. This is inconsistent with Osborne (2001) which in his study suggests that anxiety accounts for a significant portion of achievement test differences.

An examination of the anxiety variable indicates a negative correlation with outcomes on the math section of the CAHSEE. A higher level of anxiety would result in lower level performance. Also, when setting was not compared, anxiety was seen as a strong predictor of achievement for students with disabilities. However, this study compared students based on setting and found that students with disabilities in the mainstream setting indicated a higher level of anxiety than students in the self-contained setting relative to the math section of the CAHSEE. The students in the mainstream setting performed better on the math section of the exam than did the students in the self-contained setting although, the one-way analysis of covariance result indicates no statistically significant adjusted mean difference between the two groups. This implies

that anxiety felt relative to students with disabilities in the mainstream setting do not impact their performance on high stakes assessment to a significant degree.

A possible reason that students with disabilities in the mainstream setting performed better than students in self-contained setting in spite of their level of anxiety may be the result of preparedness. There is more access to the general curriculum in the mainstream setting opposed to the self-contained setting. Since students with disabilities in the mainstream setting have more access to the general curriculum, they may have felt more prepared for the academic content of the test. In other words, although students with disabilities in the mainstream setting felt the pressure of the test, it did not hinder their performance outcomes compared to students with disabilities in the self-contained setting.

Furthermore, students in the mainstream setting may have taken the test more seriously compared to the students in the self-contained setting. Students with disabilities in the mainstream setting are comparing themselves to students without disabilities and as such, may mimic the anxiety felt by other students in the setting. In the self-contained setting all students with disabilities are mimicking the behavior of other students with disabilities. Students in the self-contained setting may have felt more comfortable in the sheltered environment where most of the students are working at the same level or with the same supports as themselves.

## **Conclusions**

Stereotype threat, anxiety, and setting were factors investigated to understand their impact on CAHSEE math scores for students with disabilities. Stereotype threat has been seen as a factor in explaining low performance on standardized tests. Anxiety has been seen as a mechanism by which outcomes on performance are affected. The role

between placement and outcomes for students with disabilities continues to be a debated issue. Thus, this study investigated the impact of stereotype threat and anxiety, in regards to setting, on the CAHSEE mathematics section for first time examinees with disabilities.

On the measure of stereotype threat, this study finds that stereotype threat does not have a statistically significant impact on mean differences in the CAHSEE test scores for students with disabilities who are enrolled in mainstream or self-contained settings for math instruction when prior math ability has been controlled. Stereotype threat is a theory based upon the students' fear that they are judged based on a group association. In situations where their academic ability is being tested, the student may feel that they will confirm a negative group stereotype by which they are connected to because of group association. In other words, how they perform on a task or test is impacted by the threat of confirming the negative stereotype. In this study a negative stereotype associated with being in special education or having a disability does not significantly impact the students' performance differences on the standardized CAHSEE math section between students with disabilities in mainstream or self-contained settings.

This study also finds that anxiety for students with disabilities in mainstream and self-contained settings does not have a statistically significant impact on the mean difference in CAHSEE math test scores when prior math ability has been controlled. Anxiety is described as a response to stressful situations based on the extent to which each individual perceives the specific situation as psychologically dangerous or threatening, and is greatly influenced by each individual's past experience. This response can negatively affect the outcome on a test. In this study anxiety associated with the math section of CAHSEE, a high-stakes test does not show a statistically significant mean difference relative to the students' performance between students with disabilities in

mainstream or self-contained settings. When all participants are examined as a whole group, the multiple regression analysis implies that anxiety plays a major role in student test performance. A higher level of anxiety tends to result in a lower level of performance.

In addition, this study finds that the setting in which students with disabilities received math instruction (mainstream or self-contained) does have an impact on the mean differences in test scores on the math section of the CAHSEE. This study shows that the mean differences between the two settings are statistically significant and that students in the mainstream setting passed the CAHSEE math section on their first attempt at a higher rate than students with disabilities in the self-contained setting when prior math ability was controlled.

Stereotype threat and anxiety both are associated with the student themselves, whereas, setting is associated with the external world in which the students are educated. Stereotype threat relates to the feeling or perception the student has about himself or herself. Anxiety also relates to the student themselves, how the student is feeling at the time. Setting, however, relates to the place or institutional structure where math instruction for students with disabilities finds themselves. This study has provided evidence to suggest that when comparing students in mainstream and self-contained settings, stereotype threat and anxiety do not significantly impact the scores on the math section on the CAHSEE. This study also provided evidence to suggest that setting has a significant impact on CAHSEE math scores when comparing students with disabilities in mainstream and self-contained settings. Thus, several critical areas of investigation still need to be explored in relation to high school exit exams for students with disabilities to increase the numbers of them successfully passing on the first administration.

## **Implications**

**For research.** The findings of this study have several major implications for future research in the area of setting. Further research should examine several aspects related to high school exit exams for students with disabilities in mainstream and self-contained settings. First, future research must support and promote math achievement in the mainstream setting for students with disabilities at the high school level as found in this study. More students with disabilities are subject to passing a standardized test based on state identified levels of knowledge to receive a high school diploma. It will be important to make sure they are in the best setting for optimal achievement.

Second, future research must identify elements within the mainstream and self-contained settings at the secondary level that may contribute to increased pass rates on high school exit exams for students with disabilities. Elements related to the curriculum used to optimize the maximum knowledge needed to successfully pass the tests need to be examined. The delivery of instruction and the training provided to the instructors of students with disabilities should be examined to increase the students' math achievement and increase their success on statewide standardized tests used to grant a high school diploma.

Additionally, future research that extends this investigation by using a more in-depth instrument is recommended. A study with participants who are willing to spend more time with the instrument can include open-ended response items or interviews instead of multiple-choice items. This will allow and encourage participants to provide a fuller more meaningful answer to each question. Furthermore, a comparison could be made between students with disabilities in other suburban communities within California.

Likewise, a comparison could be made between students with disabilities in non-suburban communities within California.

Although the findings of this study do not have major implications regarding the impact of stereotype threat and anxiety on the math section of the CAHSEE for first time examinees with disabilities when comparing mainstream and self-contained settings, when students with disabilities were examined as a whole group, anxiety had a significant impact on achievement. These findings can inform decisions made by school district, site administrators, and professionals in the field of special education concerning the personal development of students with disabilities in California's suburban high schools. Further research and practice should include an investigation of ways to reduce student anxiety as a means of increasing overall student performance.

Students with disabilities in mainstream settings experienced more anxiety and less stereotype threat whereas students in self-contained settings experienced more stereotype threat and less anxiety. Future research must examine factors related to anxiety in the mainstream setting and factors related to stereotype threat need to be examined for students in the self-contained setting relative to academic performance of students with disabilities. Since stereotype threat and anxiety are student felt, and may be experienced closer to performance, it is recommended that the time between the high stakes test and the questionnaire administration be reduced to the smallest extent possible without causing undue harm to the test outcomes.

**For practice.** This study found that students with disabilities in mainstream settings for math instruction did better than students with disabilities in self-contained settings. Therefore, more students with disabilities must be placed in mainstream settings to increase their potential of higher achievement on high-stakes standardized test. This

can be facilitated by special education teachers working in collaboration with general education teachers regarding achievement for all students with disabilities.

Special education teachers can provide training and support for the various students with disabilities enrolled in the mainstream setting. Special education teachers can work collaboratively with general education teachers, counselor, and administrators in the best educational planning and placement decisions made for students with disabilities to optimize their potential. Anxiety has been shown to correlate negatively with students achievement on the CAHSEE math section, to assist with reducing the level anxiety, content on the exit exam should be introduced and practiced with the students earlier, possible 8<sup>th</sup> and 9<sup>th</sup> grades. Providing more opportunities for students with disabilities in the self-contained classrooms to interact and work with their non-disabled peers could be a strategy to reduce some of the negative stereotypes felt by students.

### **Summary**

The current study was designed to address a gap in the literature regarding low first time pass rates on exit exams for high school students with disabilities. Stereotype threat, the theoretical rationale behind the study, suggests that individuals may experience performance related stress or anxiety which can negatively affect the individual's performance when the exam is most important to the individual (Osborne, 2001; Steele, 1997). A summary and discussion of this study's findings, implications of those findings, limitations of the study, and suggestions for future research and practice were presented in this chapter.

After the examination of a sample of suburban California 10<sup>th</sup> grade students with disabilities and the relative impact of setting, stereotype threat, and anxiety on their first-

time California High School Exit Exam scores for math, suggestions were made for providing a more in-depth study of the setting in which the students receive math instruction. Future research can extend this investigation by implementing a more extensive instrument, exploring other subjective factors which may affect students with disabilities' achievement, and by examining urban communities affected by the high stakes standardized test results.

## REFERENCES

- Albrecht, S. F., & Joles, C. (2003). Accountability and access to opportunity: Mutually exclusive tenets under a high-stakes testing mandate. *Preventing School Failure, 47*(2), 86-91.
- Barton, P. E. (2004). Why does the gap persist? *Educational Leadership, 62*(3), 8-13.
- Bishop, J. H., Mane, F., & Bishop, M. (2001). How external exit exams spur achievement. *Educational Leadership, 59*(1), 58-63.
- Brookhart v. Illinois Board of Educ., 697 F.2d 179, 184 (7<sup>th</sup> Cir. 1983).
- Brown, R. P., & Day, E. A. (2006). The difference isn't Black and White: Stereotype threat and the race gap on Raven's Advanced Progressive Matrices. *Journal of Applied Psychology, 91*, 979-985.
- Brown, R. P., & Josephs, R. A. (1999). A burden of proof: Stereotype relevance and Gender differences in math performance. *Journal of Personality and Social Psychology, 76*(2), 246-257.
- Cadinu, M., Maass, A., Lombardo, M., & Frigerio, S. (2006). Stereotype threat: The moderating role of locus of control beliefs. *European Journal of Social Psychology, 36*, 183-197.
- California Department of Education (2007). California Department of Education. Retrieved March 18, 2008, from <http://www.cde.ca.gov/sp/se/lr/om112007.asp>
- California Department of Education (2008a). *Overview of the California high school exit examination (CAHSEE)*. Retrieved March 18, 2008, from <http://www.cde.ca.gov/ta/tg/hs/overview.asp>.
- California Department of Education (2008b). *The 2008-09 Academic performance index (API) Reports Information Guide*. Retrieved March 24, 2009, from <http://www.cde.ca.gov/ta/ac/ap/documents/infoguide08.pdf>
- California Department of Education (2008c). California Department of Education. Retrieved March 24, 2009, from <http://www.ced.ca.gov/nr/ne/yr08/yr08rel131.asp>
- California Department of Education (2008d). California high school exit examination (CAHSEE), Retrieved December, 2008, from <http://cahsee.cde.ca.gov/>
- California Department of Education (2009). *The 2009 Adequate yearly progress (AYP) report information guide*. Retrieved Oct. 11<sup>th</sup>, 2009, from <http://www.cde.ca.gov/ta/ac/ay/documents/infoguide09.pdf>

- Callet, V. J. (2005). High-stakes testing: Does the California high school exit exam measure up? *Language Assessment Quarterly*, 2(4), 289-307.
- Center on Education Policy. (2002). *State high school exit exams: A Baseline Report*. Washington, DC: Center on Education Policy.
- Center on Education Policy. (2004). *State high school exit exams: A maturing reform*. Washington, DC: Center on Education Policy.
- Center on Education Policy. (2005). *State high school exit exams: States try harder, but gaps persist*. Washington, DC: Center on Education Policy.
- Center on Education Policy. (2007). *State high school exit exams: Working to raise test scores*. Washington, DC: Center on Education Policy.
- Chapman, et al., v. California Department of Education, et al., No. c-01-01780 CRB JCS, U.S. Dist. (2002).
- Christenson, S. L., Decker, D. M., Triezenberg, H. L., Ysseldyke, J. E., & Reschly, A. (2007). Consequences of high-stakes assessment for students with and without disabilities. *Educational Policy*, 21(4) 662-690.
- Civil Rights Project (2005). *Confronting the Graduation Crisis in California*. Los Angeles, CA: Civil Rights Project / Proyecto Derechos Civiles at University of California at Los Angeles.
- Cohen, J. (1992). A power primer. *Psychology Bulletin*, 112(1) 155-159.
- Cole, B., Matheson, K., & Anisman, H. (2007). The moderating role of ethnic identity and social support on relations between well-being and academic performance. *Journal of Applied Social Psychology*, 37(3), 592-615.
- Cornell, D. G., Krosnick, J. A., & Chang, L. C. (2006). Students' reactions to being wrongly informed of failing a high-stakes test: The case of the Minnesota basic standards test. *Educational Policy*, 20(5), 718-751.
- Creswell, J. (2005). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research* (2nd ed.) Upper Saddle River, NJ: Merrill/Pearson.
- Debra P. v. Turlington, 644 f.2d 397(5th Cir. Unit B 1981).
- Dunn, C., Chambers, D., & Rabren, K. (2004). Variables affecting students' decisions to drop out of school. *Remedial and Special Education*, 25(5) 314-324.

- Eccles, J. S., & Wigfield, A. (1995). In the mind of the actor: The structure of adolescents' achievement task values and expectancy-related beliefs. *Personality and Social Psychology Bulletin*, *21*(3), 215-225.
- Fore III, C., Hagan-Burke, S., Burke, M. D., Boon, R. T., & Smith, S. (2008). Academic achievement and class placement in high school: Do students with learning disabilities achieve more in one class placement than another? *Education and Treatment of Children*, *31*(1) 55-72.
- Garcia, P. (2003). The use of high school exit examinations in four southwestern states. *Bilingual Research Journal*, *27*(3), 431-540.
- Ginsburg, H.P., & Baroody, A.J. (1990). *Test of Early Mathematics Ability-Second Edition*. Austin, TX: Pro-Ed.
- Good, C., Aronson, J., & Harder, J. A. (2008). Problems in the pipeline: Stereotype threat and women's achievement in high-level math courses. *Journal of Applied Developmental Psychology*, *29*, 17-28.
- Good, C., Aronson, J., & Inzlicht, M. (2003). Improving adolescents' standardized test performance: An intervention to reduce the effects of stereotype threat. *Journal of Applied Developmental Psychology*, *24*, 645-662.
- Hauser, R. M. (1997). Indicators of high school completion and dropout. In R.M. Hauser, B.V. Brown, and W.R. Prosser (Eds.), *Indicators of Children's Well-being* (pp. 152-184). New York: Russell Sage Foundation.
- Hoff, D. (2006). State urgency over exit test fades with time. *Education Week*, *25*(41), 1, 30, 32.
- Holloway, J. H. (2001). Inclusion and students with learning disabilities. *Educational Leadership*, *58*, 86-87.
- Horn, C. (2003). High-stakes testing and students: Stopping or perpetuating a cycle of Failure? *Theory Into Practice*, *24*(1), 30-41.
- Howell, K. W., Zucker, S. H., & Morehead, M. K. (1985). *Multilevel Academic Survey Test*: San Antonio, TX: The Psychological Corporation.
- Inzlicht, M., & Ben-Zeev, T. (2000). A threatening intellectual environment: Why females are susceptible to experiencing problem-solving deficits in the presence of males. *Psychological Science*, *11*, 365-371.
- Jacobson, K. G. (2000, April). *Central tensions: A critical framework for examining high school mathematics and mathematics education*. Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA.

- Johnson, D. R., Thurlow, M. L., Cosio, A., & Bremer, C. (2005). High school graduation requirements and students with disabilities. *National Center on Secondary Education and Transition Information Brief*, 4(2).
- Johnson, D. R., Thurlow, M. L., & Stout, K. E. (2007). *Revisiting graduation requirements and diploma options for youth with disabilities: A national study* (Technical Report 49). Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes.
- Kaufman, P., Alt, M. N., & Chapman, C. (2004). *Dropout rates in the United States: 2001* (NCES 2005-046). U.S. Department of Education. National Center for Education Statistics. Washington, DC: U.S. Government Printing Office.
- Kaufman, A. S., & Kaufman, N. L. (1990). *Kaufman Brief Intelligence Test*. Circle Pines, MN: American Guidance Service.
- Kaufman, A. S., & Kaufman, N. L. (1998). *Kaufman Test of Educational Achievement*. Circle Pines, MN: American Guidance Service.
- Kavale, K., & Forness, S. (2000). History, rhetoric, and reality: Analysis of the inclusion debate. *Remedial and Special Education*, 21(5), 279–296.
- Keller, J. (2007). Stereotype threat in classroom settings: The interactive effect of domain identification, task difficulty, and stereotype threat on female students' math performance. *British Journal of Educational Psychology*, 77, 323–338.
- Kellow, J. T., & Jones, B. D. (2008). The effects of stereotypes on the achievement gap: reexamining the academic performance of African American high school students. *Journal of Black Psychology*, 34(1), 94–120.
- Laird, J., DeBell, M., Kienzl, G., and Chapman, C. (2007). *Dropout rates in the United States: 2005* (NCES 2007-059). U.S. Department of Education. Washington, DC: National Center for Education Statistics. Retrieved May 16, 2008 from <http://nces.ed.gov/pubsearch>.
- Manset, G., & Washburn, S. J. (2000). Equity through accountability: Mandating minimum competency examinations for secondary students with learning disabilities. *Learning Disabilities Research & Practice*, 15, 160–167.
- Margalit, M., & Al-Yagon, M. (2002). The loneliness experience of children with learning disabilities. In B.Y.L. Wong & M. Donahue (Eds.), *The social dimensions of learning disabilities* (pp. 53–75). Hillsdale, NJ: Lawrence Erlbaum.

- Muenning, P. (2005). *The economic value of health gains associated with education interventions*. Prepared for the Equity Symposium on The Social Costs of Inadequate Education at Teachers' College, Columbia University. Retrieved April 12, 2008  
[http://www.tc.columbia.edu/i/a/3082\\_socialcostsofinadequateEducation.pdf](http://www.tc.columbia.edu/i/a/3082_socialcostsofinadequateEducation.pdf)
- Neuville, E., & Croizet, J. C. (2007). Can salience of gender identity impair math performance among 7-8 years old girls? The moderating role of task difficulty. *European Journal of Psychology of Education, 22*(3), 307-316.
- Nichols, J. D. (2003). Prediction indicators for students failing the state of Indiana high school graduation exam. *Preventing School Failure, 47*(3), 112-120.
- No Child Left Behind Act of 2001, Pub. L. No. 107-110, 115 Stat. 1425 (2002).
- O'Neil, P. T. (2001). Special education and high stakes testing for high school graduation: An analysis of current law and policy. *Journal of Law and Education, 30*(1), 185-222.
- Osborne, J. W. (2001). Testing stereotype threat: Does anxiety explain race and sex differences in achievement? *Contemporary Educational Psychology, 26*, 291-310.
- Planty, M., Hussar, W., Snyder, T., Provasnik, S., Kena, G., Dinkes, R., Kewal-Ramani, A., & Kemp, J. (2008). *The Condition of Education 2008* (NCES 2008-031). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.
- Raskind, M. H., Margalit, M., & Higgins, E. (2006). "My LD": Children's voices on the Internet. *Learning Disability Quarterly, 29*, 253-268.
- Rea, P. J., McLaughlin, V. L., & Walther-Thomas, C. (2002). Outcomes for students with learning disabilities in inclusive and pullout programs. *Exceptional Children, 68*(2), 203-222.
- Reardon, S. F., Arshan, A., Atteberry, N., & Kurlaender, M. (2009). *Effects of the california high school exit exam on student persistence, achievement, and graduation*. Stanford, CA: Working Paper Series, Institute for Research on Educational Policy and Practice, Stanford University.
- Reid, D. K., Hresko, W. P., & Hammill, D. D. (1989). *The Test of Early Reading Ability* (2nd ed.). Austin, TX: Pro-Ed.
- Roid, G. H., & Miller, L. J. (1997). *Leiter International Performance Scale-Revised*. Wood -Dale, IL: Stoelting.
- Rotter, J. B. (1966). Generalized expectancies for internal versus external control of reinforcement. *Psychological Monographs, 80* (1, Whole No. 609).

- Rouse, C. (2005). The labor market consequences of an inadequate education. Prepared for the Equity Symposium on "The Social Costs of Inadequate Education" at Teachers' College, Columbia University.
- Samuels, C. A. (2007, May 23). California students with disabilities face exit exam. *Education Week*, 26 (38), pp. 18, 23.
- Signor-Buhl, S. J., LeBlanc, M., & McDougal, J. (2006). Conducting district-wide evaluations of special education services: A case example. *Psychology in the schools*, 43(1), 109-115. Retrieved Jan. 25, 2010, online in Wiley InterScience <http://www.interscience.wiley.com>.
- Slosson, R. L. (1996). *Slosson Intelligence Test-Revised*. East Aurora, NY: Slosson Educational.
- Spencer, S. J., Steele, C. M., & Quinn, D. (1999). Stereotype threat and women's math performance. *Journal of Experimental Social Psychology*, 35, 4-28.
- Spielberger, C. D., Gorsuch, R. R., & Lushene, R. (1970). *The State-Trait Anxiety Inventory (STAI) test manual*. Palo Alto, CA: Consulting Psychologists Press.
- Steele, C. M., Spencer, S. J., & Aronson, J. (2002). Contending with group image: The psychology of stereotype and social identity threat. In M. P. Zanna (Ed.), *Advances in Experimental Social Psychology* (Vol. 34, pp. 379-440). San Diego: Academic Press, Inc.
- Steele, C. M. (1997). A threat in the air: How stereotypes shape intellectual identity and performance. *American Psychologist*, 52(6), 613-629.
- Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology*, 69(5), 797-811.
- Swanson, C. B. (2008). Cities in crisis: A special analytic report on high school graduation. Editorial Projects in Education Research Center.
- Taylor, R. P. (2006). *Factors contributing to a failure to pass and/or nonproficient score on the California high school exit exam in mathematics*. (Doctoral dissertation, Azusa Pacific University, 2006). Ann Arbor, MI: ProQuest Digital Dissertation. (UMI No. 3207577).
- Thorndike, R. L., Hagen, E. P., & Sattler, J. M. (1986). *Stanford-Binet Intelligence Scale* (4th ed.). Chicago: Riverside.

- U. S. Department of Education (2004). *The Condition of Education 2004* (NCES 2004-077). National Center for Education Statistics. Washington, DC: U.S. Government Printing Office.
- Warren, J. R., Jenkins, K. N., & Kulick, R. B. (2006). High school exit examinations and state-level completion and graduation rates, 1975 through 2002. *Educational Evaluation and Policy Analysis*, 28(2), 131-152.
- Wechsler, D. (1991a). *Wechsler Intelligence Scale for Children* (3rd. ed.). San Antonio, TX: Psychological Corporation.
- Wechsler, D. (1991b). *Wechsler Individual Achievement Test*. New York: Psychological Corporation.
- Wechsler, D. (2001). *Wechsler Individual Achievement Test* (2nd ed.). New York: Psychological Corporation.
- Wiener, J. (2004). Do peer relationships foster behavioral adjustment in children with learning disabilities? *Learning Disability Quarterly*, 27, 21-30.
- Wiener, J., & Schneider, B. H. (2002). A multisource exploration of the friendship patterns of children with and without learning disabilities. *Journal of Abnormal Child Psychology*, 30, 127-141.
- Wiesen, J. P. (1996). *Spatial Ability Test*. Newton, MA: Applied Personnel Research.
- Wilkinson, G. S. (1993). *Wide Range Achievement Test* (3rd Rev. ed.). Wilmington, DE: Jastak.
- Wise, L. L. et al. (2005). *Independent evaluation of the California high school exit examination (CAHSEE): 2005 evaluation report executive summary*. Human Resources Research Organization. Retrieved March 16, 2008 from <http://www.cde.ca.gov/ta/tg/hs/documents/year6execsum.doc>.
- Woodcock, R. W., & Johnson, M. B. (1989). *Woodcock-Johnson Psychoeducational Battery-Revised*. Itasca, IL: Riverside.
- Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). *The Woodcock-Johnson-III*. Itasca, IL: Riverside.
- Ysseldyke, J., Nelson, J. R., Christenson, S., Johnson, D. R., Dennison, A., Triezenberg, H., Sharpe, M., & Hawes, M. (2004). What we know and need to know about the consequences of high-stakes testing for students with disabilities. *Exceptional Children*, 71(1), 75-94.

- Zau, A. C., & Betts, J. R. (2008). *Predicting success, preventing failure: An investigation of the california high school exit exam*. Public Policy Institute of California. Retrieved January 31, 2008 from <http://www.ppic.org>.
- Zhang, D., Katsiyannis, A., & Kortering, L. J. (2007). Performance on exit exams by students with disabilities: A four-year analysis. *Career Development for Exceptional Individuals*, 30(1), 48-57.

**APPENDIXES**

## Appendix A- USF IRB Approval

### USF IRB APPROVAL

March 26, 2010

Dear Ms. Fields:

The Institutional Review Board for the Protection of Human Subjects (IRBPHS) at the University of San Francisco (USF) has reviewed your modification request for human subjects approval regarding your study.

Your modification application has been approved by the committee (IRBPHS #09-092). Please note the following:

1. Approval expires twelve (12) months from the dated noted above. At that time, if you are still in collecting data from human subjects, you must file a renewal application.
2. Any modifications to the research protocol or changes in instrumentation (including wording of items) must be communicated to the IRBPHS. Re-submission of an application may be required at that time.
3. Any adverse reactions or complications on the part of participants must be reported (in writing) to the IRBPHS within ten (10) working days.

If you have any questions, please contact the IRBPHS at (415) 422-6091.

On behalf of the IRBPHS committee, I wish you much success in your research.

Sincerely,

Terence Patterson, EdD, ABPP  
Chair, Institutional Review Board for the Protection of Human Subjects

-----  
IRBPHS – University of San Francisco  
Counseling Psychology Department  
Education Building – Room 017  
2130 Fulton Street  
San Francisco, CA 94117-1080  
(415) 422-6091 (Message)  
(415) 422-5528 (Fax)  
[irbphs@usfca.edu](mailto:irbphs@usfca.edu)

-----  
<http://www.usfca.edu/humansubjects/>

**Appendix B- Student Questionnaire Coversheet**

**Student ID #** \_\_\_\_\_  
**Math Course/Teacher**

---

**California High School Exit Exam and 10<sup>th</sup> Grade****High School Students**

## Introduction

Thank you for taking time to complete this questionnaire. The purpose of the questionnaire is to get your opinions and thoughts regarding the California High School Exit Exam (CAHSEE). As you may know passage of the CAHSEE is a requirement for most students who want to graduate from high school with a diploma. Therefore, please answer each question as accurately as possible.

***THANK YOU AGAIN!***

For questions regarding the questionnaire or study, contact Tracy Fields

## Appendix C- Student Anxiety Questionnaire

### Student Questionnaire

**INSTRUCTIONS:** Read each statement and check inside the box which best indicates your level of agreement. Check the box which seems to best describe your present feelings.

<b>“When taking the math section of the CAHSEE, I felt...”</b>	<b>Strongly Disagree</b>	<b>Disagree</b>	<b>Agree</b>	<b>Strongly Agree</b>
1. Calm				
2. Secure				
3. Tense				
4. Strained				
5. At ease				
6. Upset				
7. Satisfied				
8. Frightened				
9. Comfortable				
10. Confident				
11. Nervous				
12. Jumpy				
13. Unsure				
14. Relaxed				
15. Worried				
16. Confused				
17. Steady				

### Appendix D- Student Stereotype Threat Questionnaire

**INSTRUCTIONS:** Read each statement and check inside the box which best indicates your level of agreement. Check the box which seems to best describe your present feelings.

	Strongly Disagree	Disagree	Agree	Strongly Agree
1. If I failed the CAHSEE, people will think I have less ability because I have a disability.				
2. Because of my disability, people will look down on me if I failed the CAHSEE.				
3. The CAHSEE may have been easier for other people.				
4. Because I have a disability, taking the test made me question my knowledge of math.				
5. Because I have a disability, I was concerned about whether I had enough math skills to pass the CAHSEE.				
6. Because I have a disability, I feel less confident about my math skills when I'm in my math class.				
7. Because I have a disability, I am looked upon negatively in my math class.				
8. People expect me to do poorly on the CAHSEE because of my disability.				
9. Others believe that my disability determines how I did on the CAHSEE.				

## Appendix E- Validity Panel Questionnaire

### VALIDITY PANEL QUESTIONNAIRE

**DIRECTIONS:** Please take a few minutes to tell me about yourself by answering the questions below.

1. Name:

2. Official Title:

3. Highest Level of Education (Circle One):

Doctorate

Master's Degree

Bachelor's Degree

4. Number of years in the education profession:

5. Number of years in special education or working with special education:

6. Do you teach currently, if so list the courses or content area(s) and level you have taught within the past 12 months? \_\_\_\_\_

---



---

7. If you are a K-12 public educator are you considered highly qualified by the state of California?      Yes      No

8. Do you currently hold a *valid* K-12 teaching credential? Yes      No

9. Please list any other type of *valid* credential you currently hold:

10. Please feel free to add any additional information you about yourself that may be helpful to the study.

**THANK YOU!!**

## Appendix F- Validity Panel Demographics

### Validity Panel Demographic Data

	#1	#2	#3	#4	#5	#6	#7
Title	Assistant Principal	SPED Teacher on Special Assignment	Special Education Teacher	Program Specialist	Program Specialist	Special Education Teacher	Career Transition Specialist/Career Counseling and Guidance
Highest Level of Education	Masters	Masters	Bachelors	Masters	Bachelor of Arts + 90 units	Doctorate	Masters
# of years in education	32 years	15 years	20 years	28 years	45 years	20 years	34 years
# of years in special education	2 years	15 years	20 years	20 years	45 years	8 years	28 years
content area(s) teaching within past 12 months	NA	Special Education Transition	SDC Math and Science; CAHSEE prep. for SPED students	NA	NA	SDC English, Social Studies, and Math	Secondary Level Pre-employment, skills training, and Independent living skills
Highly Qualified California K-12 public educator	NA	Yes	Yes	NA	NA	Yes	Yes
Current K-12 teaching credential	No	Yes	Yes	Yes	Yes	Yes	Yes
Other valid credentials	PPS; Administrative Services	Educational Specialist	Educational Specialist; Single-Subject Science	Single-Subject English; Resource Specialist; Learning Handicapped	K-8 GE; Resource Specialist; Exceptional Children K-12; CLAD Certificate	Single/Subject Social Studies; Moderate/Severe SPED.	Severely Handicapped Specialist Credential
Additional information					taught all levels grades 1-12. I was in the classroom three years ago teaching SPED 6-8.		Career Transition Certificate; CCLAD Certificate

## Appendix G- Validity Panel Packet

Dear Extremely Busy Educator,

Thank you in advance for taking a few minutes to read this letter. The purpose of this correspondence is to invite you to be part of a **panel of experts** in a research project I am conducting which involve secondary students with disabilities in Northern California's suburban schools. This study has been approved by the university's Institute for the Protection of Human Subjects and aims to survey approximately 60-80 secondary students with disabilities on several measures centered on the California High School Exit Exam. Without your input I am unable to move forward with this research.

Attached is a draft of the survey I intend to use with secondary students with disabilities as part of my study. What I need from you is a critique of the questionnaire and your written comments regarding the format, length, and appropriateness of the questions on the survey. Please answer the following questions about the survey. The critique should take no longer than 30 minutes to complete. Should you have any questions or would like information regarding the findings of my study, please do not hesitate to contact me via email.

Thank you again for your time and I look forward to your comments and critique of my survey instrument.

Sincerely,

Tracy Fields

Below is a definition of stereotype threat which this proposed study is grounded. A questionnaire will be administered to a sample of 10<sup>th</sup> grade students with disabilities in public suburban high schools in Northern California's Bay Area. Participants will read the questions and complete the questionnaire independently within a group or individual setting administration.

**Stereotype Threat:** Stereotype threat is the belief or threat of confirming a negative stereotype about a group to which one belongs. Individuals in stereotype-threatened conditions consistently underperform compared to their non-threatened counterparts. Anyone with a group identity to which a negative stereotype exists can be affected. Stereotype threat describes the possibility of confirming a negative stereotype, *not* the actual experience. Thus, the stereotype does not have to be real. However, if the threat is experienced in the midst of a presentation or test, for example, the emotional reaction it causes may interfere with a students' performance.

**Directions:** Please respond to the questions below regarding the attached survey.

1. Please comment on the appearance of the survey:
2. Does the questionnaire appear to measure the constructs of the theory to which it is grounded?
3. Please comment on the language used on the survey in terms of clarity (refer to specific item numbers when possible).
4. Please write any additional information, comments or recommendations you have regarding the survey attached.

The primary purpose of this proposed project is to gain an initial line of inquiry into the impact of setting, anxiety, and stereotype threat on the math section passage rates

of the California High School Exit Exam (CAHSEE) for students with disabilities who are in mainstream or self-contained settings for math instruction. Given the aforementioned objective, please comment on the extent to which the questions address the setting, and the CAHSEE:

1. Is there evidence of content relevance in the survey questions? Please refer to specific survey item numbers in your response.

2. Have I covered everything or is something missing in terms of the subject matter and/or theoretical rationale?

3. Does the content contained in the survey appear to have generalizability to the proposed population?

4. Are the constructs of stereotype threat –anxiety, evaluation apprehension, and self-efficacy- represented in the survey questions?

## Appendix H- Validity Panel Comments

<b>FACE VALIDITY</b>
<p>#1- It looks non-threatening and easy to complete. It appears to measure the constructs of the theory. Question 1 might be difficult for a 10<sup>th</sup> grader to answer because of double negative. Might elicit an incorrect answer. Question 5 is vague. Question 8 ? Question 14-I wouldn't know how to answer it. Question 15- By the time you get to the end you might forget the "Because I have a disability" part. Consider putting disability in question.</p>
<p>#2-Format of the survey is well done. I would give the survey a name. Remove the statement "Because I have a disability" and incorporate into directions. It does not fit each statement below well. The language of the statements is clear and simple to understand what is being asked. One concern/question I had was whether to begin the survey with negative statements, which may make the student more resistant or oppositional to completing the rest of the survey. Also the opening phrase so as not lead well into each of the statements in the survey. Also, some students may not own their disability and may not associate with the opening phrase.</p>
<p>#3-The survey seems open, friendly, and clear in presentation. The instructions seem easy to follow, the font large enough to read. I liked the spacing between 3 and 4, 5 and 6, and would like a little more space between 7 and 8, 13 and 14. Is the theory regarding the stereotyping of students taking the math section of the CAHSEE? If I got that right, then the survey would seem to address that question. Are you supposed to preface the survey question with "Because I have a disability"? If so, some of the questions seem a little different. Number 9- Because I have a disability, others view my ability positively (?).</p>
<p>#4-Instructions are quite clear, limited number of responses, choices for individual perceptions, what stands out..."Because ..." doesn't that just beg the question? When I read, actually, before I read the page what just jumped out was..."Because I have a disability" which in itself stereotypes. If one buys into statistics, yes indeed! (Again, what it is you seek just is too obvious..."Because..." It seems the end result is to clarify/validate, confidence, pre-test assumed failure, " not caring about other's belief. No suggestions.</p>
<p>#5-The survey appeared to me to be easy to read and understand. It looked good. Statements appeared clear to me and were easily understood. Number 15-you may want to state this differently, so that it makes a complete sentence with the beginning of the sentence (just a matter of <u>format</u>).</p>
<p>#6- It looks fine to me. Not too complicated, nice and simple. It looks okay, no big words.</p>
<p>#7- The survey directions are clear. The print is a good size and easy to read. Why isn't there a period at the end of each survey statement? Setting, stereotype threat-anxiety, apprehension, self-efficacy) Yes, as far as I can tell the questions appear to measure the constructs of the theory to which it is grounded. #15 "Because I have a disability ..." What? My disability does not affect people's view of my ability. People's view of my ability is not affected. For consistency, the presentation of the survey should have a script. I believe it should be presented to each participant orally and visually in the same scripted manner to make sure the survey measures what it claims to measure. I know you want the results to be accurately applied and interpreted.</p>

<b>CONTENT VALIDITY</b>
#1- Yes, there is evidence of content relevance on Stereotype measure: setting #8, 11, 13; Anxiety #2, 4, 5, 7; Stereotype #1, 5, 11, 12, 13, 14, 15. Covered everything. Yes, it appears to have generalizability.
#2- Overall, the surveys seems to pinpoint one's feelings about their math ability, taking the CAHSEE and having a disability. Yes, generalizability.
#3- The survey Appendix B seems to cover content relevance well, especially the questions that ask a similar question worded slightly differently to check responses for validity. A statement could be "I know I will do poorly on the CAHSEE" as a baseline. Yes, you seem to have enough of a student's sense of paranoia and persecution, with just a touch of grandiosity.
#4- Yes, there is evidence of content relevance. You covered stereotype threat for CAHSEE and subject. Oh yes, immediate generalizability.
#5- Yes, the statements appear relevant to me. The survey appears to cover both positive and negative perceptions. Yes, there is generalizability.
#6- Evidence of content relevance, I think so. The questions seem to refer to the CAHSEE. It seems to be okay (you have covered everything). Yes, there is generalizability.
#7- Evidence of content relevance, setting mainstream, self-contained, and CAHSEE math. Appendix B survey statements #9 and #15 ability in general or math ability. In the survey statements you do cover setting and level of agreement regarding the three stereotype threats- anxiety, apprehension. For generalization to 10 <sup>th</sup> graders with disabilities in public suburban high schools in Northern California, you would need more than 60-80 participants. Yes, 10 <sup>th</sup> graders would be able to relate to, and answer the survey statements regarding their feelings about the statements.

<b>CONSTRUCT VALIDITY</b>
#1- See previous page
#2- Yes, the initial opening phrase to each statement hones in on the fact the student has a disability and that is a focus that can cause anxiety as soon as one begins the survey. The first two statements seem focused on the negative perceptions of having a disability which can affect the next responses. The questions then seem to address one's confidence in their abilities despite having a disability.
#3- There seems to be a number of questions addressing the student's anxiety with the math section of CAHSEE, and math in general. Would statements regarding their anxiety of academic competence, in general, be appropriate? There are a number of statements that address the student's concerns about being judged by staff and/or peers. There are some questions addressing self-efficacy, but not as much as the statements for anxiety and evaluation apprehension.
#4- Yes, certainly, yes
#5- Appears to be addressed.
#6- Definitely, yes, yes.
#7- Are these (#3,6,7,8) the statements for anxiety? Are these (#1,2,4,5,10,11,12,13,14) the statements for evaluation apprehension? Are these (# 9,15) the statements for self-efficacy?

## Appendix I- School District Consent

### PREMISSION LETTER FROM CONSENTING SCHOOL DISTRICTS OR HIGH SCHOOLS

# CASTRO VALLEY HIGH SCHOOL

*Castro Valley High School, with a commitment to excellence, empowers the individual with knowledge, a passion for life-long learning, and a respect for the principles of our diverse democratic society.*

10400 Santa Maria Avenue • Castro Valley, California 94546 •

December 7<sup>th</sup>, 2009

Institutional Review for the Protection of  
Human Subjects University of San Francisco  
2130 Fulton Street  
San Francisco, CA 94117

Dear Members of the Committee:

On behalf of the Castro Valley High School, I am writing to formally indicate our awareness of the research proposed by Mrs. Tracy M. Fields, a doctoral student at USF. We are aware that Mrs. Fields intends to conduct her research by administering a written questionnaire to our students receiving special education services, and by conducting a review of their educational records.

I am responsible for the faculty and staff of my school. I give Mrs. Fields permission to conduct her research at our school site. If you have any questions or concerns, please feel free to contact my office at (510) 537-5910.

Sincerely,



Pete Alvarez  
Principal, Castro Valley High School



All Dublin Students Will  
Become Lifelong Learners

## DUBLIN SCHOOLS

DUBLIN UNIFIED SCHOOL DISTRICT

7471 Larkdale Avenue, Dublin, CA 94568-1599 | 925-828-2551 | FAX 925-829-6532

March 25, 2010

Institutional Review for the Protection of Human Subjects

University of San Francisco

2130 Fulton Street

San Francisco, CA 94117

Dear Members of the Committee:

On behalf of the Dublin Unified School District, I am writing to formally indicate our awareness of the research proposed by Mrs. Tracy M. Fields, a doctoral student at USF. I give Mrs. Fields permission to conduct her research within our district contingent on approval from the USF Institutional Review Board. If you have any questions or concerns, please feel free to contact my office at (925) 828-2551.

Sincerely,

A handwritten signature in black ink, appearing to read "Pain Lear".

Pain Lear  
Director of Educational Services  
Dublin Unified School District

# New Haven Unified School District

UNION CITY • SOUTH HAYWARD •  
34200 Alvarado-Niles Road • UNION CITY • CA 94587

March 3, 2010

Institutional Review for the  
Protection of Human Subjects  
University of San Francisco  
2130 Fulton Street  
San Francisco, CA 94117

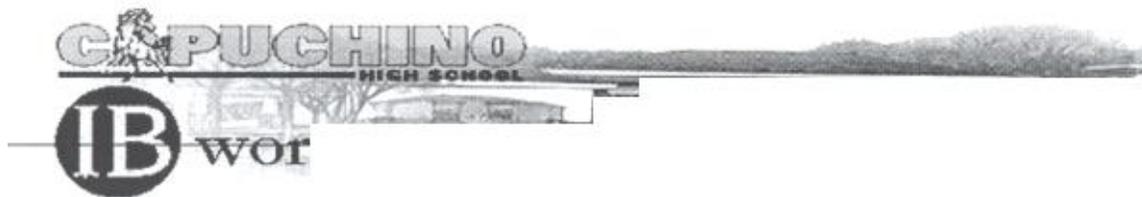
Dear Members of the Committee:

On behalf of the New Haven Unified School District, I am writing to formally indicate our awareness of the research proposed by Mrs. Tracy M. Fields, a doctoral student at USE. We are aware that Mrs. Fields intends to conduct her research by administering a written questionnaire to our students receiving special education services, and by conducting a review of their educational records.

I am responsible for the faculty and staff of my school. I give Mrs. Fields permission to conduct her research at our school site. If you have any questions or concerns, please feel free to contact my office at (510)489-4141.

Sincerely,

Tasha L. Dean, Ed. D.  
Coordinator of Special Services



January 21, 2010  
Institutional Review for the Protection of Human Subjects  
University of San Francisco  
2130 Fulton Street  
San Francisco, CA 94117  
Dear Members of the Committee:

On behalf of the Capuchino High School, I am writing to formally indicate our awareness of the research proposed by Mrs. Tracy M. Fields, a doctoral student at USF. We are aware that Mrs. Fields intends to conduct her research by administering a written questionnaire to our students receiving special education services, and by conducting a review of their educational records.

I am responsible for the faculty and staff of my school. I give Mrs. Fields permission to conduct her research at our school site. If you have any questions or concerns, please feel free to contact my office at (650) 558-2799.

Sincerely,

A handwritten signature in cursive script, which appears to read 'Shamar Shanks', is written in black ink.

Shamar Shanks  
Principal  
Capuchino High School

## Appendix J- Parental Consent Packet

### PARENT COVER LETTER

Dear Sir or Madam:

My name is Tracy M. Fields. I am a Resource Specialist and a doctoral candidate in the Department of Learning & Instruction at the University of San Francisco. I am writing to you to request your consent for your child to participate in a research study that I am doing to investigate the relationship between student performances on the mathematics section of the California High School Exit Exam (CAHSEE) among students with disabilities who are enrolled in a mainstream or self-contained setting for math. I want to know whether 10<sup>th</sup> grade students with disabilities placed in mainstream or self-contained settings for mathematics instruction differ in relation to a perceived stereotype, anxiety, and their passage or failure status on the math section of the CAHSEE. The intent of this study is to improve in the pass rates of students with disabilities on the CAHSEE.

If you allow your child to participate, after the administration of the mathematic section of the CAHSEE in the spring, your child will be asked to complete a questionnaire. The questionnaire is designed to measure anxiety and stereotype threat that your child may have experienced during the test. The questionnaire consists of 26 questions and will take approximately 15-20 minutes to complete.

Rest assured that I will also take steps to maintain confidentiality of your child's records by keeping all data materials and academic records in a locked filing cabinet at my home. All records will remain confidential and your child's participation or nonparticipation will in no way negatively affect the quality of education your child receives, or the quality of services you receive as a parent of a child attending the researcher's school. There will be no cost to your child for participating. At your own request, I will provide you with a copy of the completed study at no cost. There will be no payment available to you for your child's participation; however, your child may be **eligible** to participate in a raffle drawing to receive a **\$25.00 gift card**.

Participation in research is voluntary. You are free to decline your child's participation in this study, or withdraw from it at any point. The school is aware of this study, but is not requiring that your child participate in this research and your decision as to whether or not to participate will have no influence on the quality of education your child will receive, nor will your child's participation or non-participation influence future interactions between him/herself and school personnel.

Thank you for your consent and assistance with this study. If you have any questions please feel free to contact me by phone or by email or by regular mail.

Sincerely,

**Tracy M. Fields**  
USF Doctoral Candidate

## CONSENT TO BE A RESEARCH SUBJECT

### Purpose and Background

Mrs. Tracy M. Fields, a doctoral candidate in the School of Education at the University of San Francisco is doing a study to investigate the relationship between student performances on the mathematics section of the California High School Exit Exam (CAHSEE) among students with disabilities who are enrolled in a mainstream or self-contained setting for math.

My child is being asked to participate because he/she is classified as students with disabilities and identified as having an Individualized Education Plan (IEP) by the district. In addition, my child is enrolled in his or her sophomore year of high school and will be taking the CAHSEE for the first time in the spring administration of the test.

### Procedures

If I allow my child to be a participant in this study, the following may happen:

1. My child will take the mathematics section of the CAHSEE administered by the school district during the spring semester of their sophomore year in high school.
2. Mrs. Fields will have access to my child's relevant documents/educational records (which will remain confidential).
3. My child will complete a questionnaire about the CAHSEE.

### Risks and/or Discomforts

1. It is possible that some of the questions on the questionnaire may make my child feel uncomfortable, but he/she is free to decline to answer any questions or to stop participation at any time.
2. Participation in research may mean a loss of confidentiality. Study records will be kept as confidential as is possible. No individual identities will be used in any reports or publications resulting from the study. Study information will be coded and kept in locked files away from the school site at all times. Only study personnel will have access to the files.

### Benefits

There will be no direct benefit to me from participating in this study. The goal is to increase future pass rates on the CAHSEE for students with disabilities.

### Costs/Financial Considerations

There will be no financial costs to me or to my child as a result of taking part in this study.

### Payment/Reimbursement

There will be no payment for my child's participation in this study. However, my child may be eligible to participate in a raffle drawing to receive a \$25.00 gift card.

### Questions

I have talked to Mrs. Fields about this study and have had my questions answered. If I have further questions about the study, I may call her or email her.

If I have any questions or comments about my child's participation in this study, I should first talk with the researcher. If for some reason I do not wish to do this, I may contact the IRBPHS, which is concerned with protection of volunteers in research projects. I may reach the IRBPHS office by calling (415) 422-6091 and leaving a voicemail message, by e-mailing IRBPHS@usfca.edu, or by writing to the IRBPHS, Department of Psychology, University of San Francisco, 2130 Fulton Street, San Francisco, CA 94117-1080.

### Consent

I have been given a copy of the "Research Subject's Bill of Rights" and I have been given a copy of this consent form to keep. PARTICIPATION IN RESEARCH IS VOLUNTARY. I am free to decline to be in this study, or to withdraw my child from it at any point. My decision as to whether or not to participate in this study will have no influence on my present or future status as a student at DHS.

My signature below indicates that I agree to participate in this study.

---

Student's Name (Print)	Signature of Parent/Guardian	Date of Signature
------------------------	------------------------------	-------------------

---

Signature of Person Obtaining Consent	Date of Signature
---------------------------------------	-------------------

(PLEASE KEEP ONE COPY OF THE CONSENT FORM FOR YOUR RECORDS)

## Appendix K- Subjects' Bill of Rights

### RESEARCH SUBJECTS' BILL OF RIGHTS

The rights below are the rights of every person who is asked to be in a research study. As a research subject, I have the following rights:

- (1) To be told what the study is trying to find out;
- (2) To be told what will happen to me and whether any of the procedures, drugs, or devices are different from what would be used in standard practice;
- (3) To be told about the frequent and/or important risks, side effects, or discomforts of the things that will happen to me for research purposes;
- (4) To be told if I can expect any benefit from participating, and, if so, what the benefit might be;
- (5) To be told of the other choices I have and how they may be better or worse than being in the study;
- (6) To be allowed to ask any questions concerning the study both before agreeing to be involved and during the course of the study;
- (7) To be told what sort of medical or psychological treatment is available if any complications arise;
- (8) To refuse to participate at all or to change my mind about participation after the study is started; If I were to make such a decision, it will not affect my right to receive the care or privileges I would receive if I were not in the study;
- (9) To receive a copy of the signed and dated consent form; and
- (10) To be free of pressure when considering whether I wish to agree to be in the study. If I have questions, I should ask the researcher or the research assistant. In addition, I may contact the Institutional Review Board for the Protection of Human Subjects (IRBPHS), which is concerned with protection of volunteers in research projects. I may reach the IRBPHS by calling (415) 422-6091, by electronic mail at [IRBPHS@usfca.edu](mailto:IRBPHS@usfca.edu), or by writing the USF IRBPHS, Department of Counseling Psychology; Education Building, 2130 Fulton Street, San Francisco, CA 94117-1080.

### Appendix L- Student Data Coding Sheet

**Participants ID #** \_\_\_\_\_

**First Name** \_\_\_\_\_ **MI** \_\_\_\_\_ **Last Name** \_\_\_\_\_

**Gender:** 1) Male      2) Female      **SES:** 1) Free      2) Reduced      3) Full

**Ethnicity:**

1) White      2) African American      3) Latino      4) Asian

5) Pacific Islander      6) Alaska/Native American      7) Multi-Ethnic (Other)

**Primary Disability:**

1) LD      2) ED      3) Speech & Language      4) Visual Impairment

5) Hearing Impairment      6) Autism/Autistic-Like      7) Other

**ELL:** 1) No      2) Yes

**Math CST Score:** 8<sup>th</sup> grade \_\_\_\_\_ 9<sup>th</sup> grade \_\_\_\_\_

**Algebra 1 Course Grade:** 1<sup>st</sup>. Sem \_\_\_\_\_ 2<sup>nd</sup>. Sem \_\_\_\_\_

**Algebra 1 Course Setting:** GE \_\_\_\_\_ SPED \_\_\_\_\_

**Year Algebra 1 Course Taken:** 1) 8<sup>th</sup> grade      2) 9<sup>th</sup> grade      3) 10<sup>th</sup> grade

**9<sup>th</sup> Grade Overall GPA:** Fall \_\_\_\_\_ Spring \_\_\_\_\_

**Math CAHSEE Score:** \_\_\_\_\_ **CAHSEE Status:** Pass      Fail