

**AN ANALYSIS OF PERFORMANCE OF ESL STUDENTS ON VARIOUS
SOCIAL STUDIES OBJECTIVES AND TEST ITEMS ON THE TEXAS
ASSESSMENT OF KNOWLEDGE AND SKILLS (TAKS) TEST**

A Dissertation

by

AHLAM MUSA

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Approved by:

Chair of Committee,	Lynn M. Burlbaw
Committee Members,	Janet Hammer
	L. Quentin Dixon
	Derya Akleman
Head of Department,	Yeping Li

May 2013

Major Subject: Curriculum and Instruction

Copyright 2013 Ahlam Musa

ABSTRACT

The content area of social studies has not been given as much attention as its counterparts despite its inclusion in the state wide TAKS exam in Texas. All students in grades 8, 10, and 11 are required to take the social studies portion, and must pass it in order to graduate. The consequences of such a test and the importance of social studies in the everyday lives of students make this content area important. In addition, since ESL students constitute a considerable segment of the student population in our schools nowadays, it is essential that educators understand the challenges these students face and the experiences they go through in such content areas and tests; thus, their performance in the TAKS exam is examined in this study. The purpose of this study was to analyze the performance of ESL students in the social studies TAKS exams in grades 8 and 11 in relation to content (represented by the five objectives of the test), and question item format (limited to two text-enriched question types: excerpt-based, and bullet-point questions). The study also sought to find the effects of demographic variables on student performance. The results of the study can further familiarize educators with the challenges ESL students face in social studies to be able to assist them succeed in their classrooms and on standardized tests.

This study utilized secondary data analysis with a set of data provided from the Texas Education Agency. The data included information of the students who took the Spring 2003, 2006, and 2009 social studies TAKS exams, which allowed the researcher

to conduct longitudinal analysis to further support the results. Overall, the study was non-experimental and descriptive in nature.

A statistical significant difference was found between the percentages of correct answers by objective. ESL students performed better on objectives 3 (economics) and 5 (social studies skills), while lagging behind in objective 1 (history). The results were also confirmed in the longitudinal study that showed that although ESL students' performance increased in all five objectives, there was statistically significant difference in the performance of ESL students among the objectives. In relation to question type, ESL students chose more correct answers for the bullet-point questions than they did for the excerpt-based questions with statistical significance. Longitudinal analysis of excerpt-based questions only showed that ESL students who took the exam in 11th grade three years after taking it in 8th grade, achieved higher. Finally, the results of the study show that the three demographic variables of gender, ethnicity, and socio-economic status had an effect on the performance of ESL students. Overall, males scored higher than females, economically non-disadvantaged students outperformed disadvantaged students, and Hispanic ESL students had the lowest scores of all the ethnic groups.

DEDICATION

To Lujain and Layan. I love you.

To all the Daraghmehs and the Alqattawis; I hope this makes you proud of me and will inspire and motivate you to walk my path.

ACKNOWLEDGEMENTS

My first and ultimate thanks go to the family I grew with my ever-supporting husband, Mohammad, and to the family in which I grew (Talaat, Yazmin, Haitham, Venus, and Ashtar). Thank you for teaching me all that's important in life to help create who I am today. Your endless unconditional love, support, and belief made me a better person, and helped me realize my dream. Thank you for all the encouragement and for every single word and effort to push me forward.

I would like to thank my committee chair, Dr. Burlbaw for supporting me and believing in my abilities. Without your magic powers every now and then to get things done, this would have not been possible. I would also like to give special thanks to Dr. Dixon who gave me the opportunity to work with her and publish professionally. Thank you for all the feedback and support. Dr. Hammer, thank you for having a smile every time I saw you and for always surrounding everybody with positive energy. I would also like to thank my committee member, Dr. Akleman for putting up with my numerous statistical questions.

Thanks also go to my friends who had to put up with me and my complaints, my demands, and endless journey to show me that this could be done. Thank you for always being there to support and encourage me.

Special thanks go to my husband Mohammad and to my beautiful daughters (Lulu & Layyunni) for being the perfect family and for their patience and love. Thank you for making me a stronger person. I love you.

NOMENCLATURE

AEIS	Academic Excellence Indicator System
ESL	English as a Second Language
LEP	Limited English Proficiency
NCLB	No Child Left Behind
SES	Social Economic Status
TAAS	Texas Assessment of Academic Skills
TAKS	Texas Assessment of Knowledge and Skills
TEA	Texas Education Agency
TEKS	Texas Essential Knowledge and Skills

TABLE OF CONTENTS

	Page
ABSTRACT	ii
DEDICATION	iv
ACKNOWLEDGEMENTS	v
NOMENCLATURE	vi
TABLE OF CONTENTS	vii
LIST OF FIGURES	ix
LIST OF TABLES	x
CHAPTER I INTRODUCTION	1
Statement of the Problem	2
Purpose Statement	4
Research Questions	5
Limitations	5
Delimits	7
Definitions	8
Organization of Chapters	11
CHAPTER II LITERATURE REVIEW	13
Accountability and Standardized Testing	14
The Texas Assessment of Knowledge and Skills (TAKS)	17
Social Studies and TAKS	20
Content Knowledge versus Skill	25
ESL Students and Social Studies TAKS	27
Social Studies TAKS Test Items and Enumeration Cues	34
The Influence of Demographics on Academic Achievement	38
Gender and Achievement	40
Ethnicity and Achievement	41
Social Economic Status (SES) and Achievement	42
Summary	44

CHAPTER III DESIGN OF THE STUDY	46
Theoretical Framework	46
Subjects of Study.....	48
Data	50
Instrumentation.....	54
Analytic Framework.....	59
Participants	59
Organization of TAKS Test Question Items	64
Analytic Process	67
CHAPTER IV PRESENTATION AND ANALYSIS OF DATA	72
Research Question 1: Analysis of Performance by Objective	72
Research Question 2: Systematic Analysis of Performance by Objective.....	76
Research Question 3: Analysis of Performance by Question Format.....	80
Research Question 4: Systematic Analysis of Performance on Excerpt-Based Question Type	82
Research Question 5: Analysis of Performance by Demographic Effects.....	83
Summary of Findings.....	87
CHAPTER V SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS.....	89
Purpose of the Study	89
Conclusions and Discussion of Research Questions 1& 2: Analysis by Objective....	90
Conclusions and Discussion of Research Questions 3 & 4: Analysis by Question Type.....	94
Conclusions and Discussion of Research Question 5: Analysis of Demographic Effects.....	97
General Implications for Theory and Practice	102
Discussion of Limitations and Future Recommendations.....	107
REFERENCES.....	111

LIST OF FIGURES

	Page
Figure 1 An Image of the Spring 2006 Grade 8 Social Studies TAKS Test Data Received from the Texas Education Agency.....	52
Figure 2 Sample of Separating Student Responses and Matching Objectives to Question Items (Spring 2006, 8 th grade questions).....	54
Figure 3 Sample Excerpt-Based Question item from the 8 th grade Social studies Released TAKS Exam for the Year 2006.....	65
Figure 4 Sample Bullet-Point Question Items from the 11 th Grade Social Studies Released TAKS Exam for the Year 2009.....	66
Figure 5 Comparison Between 8 th and 11 th Grade Social Studies TAKS Test Percentages of Correct Responses per Objective.	79

LIST OF TABLES

	Page
Table 1 Social Studies Objectives in the TAKS Test, Student Expectations, and Sample Topics	23
Table 2 The Number of Questions for Each Social Studies Objective.....	24
Table 3 TAKS Passing Rates (Percentages) For All TAKS Tests in 2003, 2006, and 2009 by Ethnicity.....	40
Table 4 Internal Consistency (KR20 Reliability) for Grades 8 Social Studies Exam in 2003, 2006, and 2009.	56
Table 5 Internal Consistency (KR20 Reliability) for Grades 11 Social Studies Exam in 2003, 2006, and 2009.	56
Table 6 KR20 Reliability Estimates for the Social Studies TAKS Test in 2003, 2006, and 2009.....	57
Table 7 Gender, Ethnicity, and SES of Participants in the ESL Program Completing the Social Studies TAKS Test in Total.....	61
Table 8 Gender, Ethnicity, and ESL Status of 8 th Grade Participants in the ESL Program Completing the Social Studies TAKS Test in Spring 2003 and Spring 2006.....	62
Table 9 Gender, Ethnicity, and ESL Status of 11 th Grade Participants in the ESL Program Completing the Social Studies TAKS Test in Spring 2006 and Spring 2009.....	63
Table 10 Descriptive Statistics for the Performance of ESL students in the Social Studies TAKS in the Five Objectives.....	73
Table 11 Summary of One-Way ANOVA to Analyze the Performance of ESL Students in Social Studies TAKS Based on Test Objectives.	74
Table 12 Summary of the Welch Test to Analyze the Performance of ESL Students in Social Studies TAKS Based on Test Objectives.	75
Table 13 Descriptive Statistics for the 8 th and 11 th Grade Social Studies TAKS Test Percentages of Correct Responses per Objective.	79

Table 14 Simultaneous Multiple Regression Analysis Summary for Gender,
Ethnicity, and SES Predicting % Raw Score in the 2003, 2006, and 2009
Social Studies TAKS Tests.....86

Table 15 TAKS Passing Rates on the Social Studies TAKS Test by Gender for the
Years 2003, 2006, and 2009.98

CHAPTER I

INTRODUCTION

The number of Limited English Proficiency (LEP) students in Texas public schools has increased by 45.8 percent since 2000 (Texas Education Agency, 2011a). In the state of Texas, a total of 797,683 students receive bilingual or English as a Second Language (ESL) services (TEA, 2011a). And as a consequence of the No Child Left Behind Act (NCLB), these students are required to take the state administered Texas Assessment of Knowledge and Skills tests, better known as the TAKS, three years after arriving to the United States. These tests, which are needed to graduate, exert much pressure on LEP students in the secondary level as these students are limited in terms of time, academic linguistic knowledge of English, and subject matter background (Duff, 2001). As a result, they may be performing and achieving at low rates (Duff, 2001). Accordingly, students, teachers, and schools are trying to find ways to help this student segment (Case, 2006; Duff, 2001; Segall & Gaudelli, 2007). Of the content areas tested on the TAKS, social studies is the most recent to be added, and the most alien to LEP students in terms of content (Chamot & O'Malley 1994; Cho & Reich, 2008; Coltrane, 2002; Duff, 2001; Eslami & Plett, 2008; Reilly, 1988; Seixas 1993). In an attempt to address the challenges ESL students face with social studies, this study aims at identifying specific areas of weakness among ESL students through an analysis of their performance on the TAKS social studies tests which test knowledge of the five objectives designated by TAKS in this content area. In addition, an analysis of the text

organization in certain social studies TAKS question items is done to further understand these students' performance on the test.

Statement of the Problem

Despite the fact that the content area of social studies is not required to be assessed by the NCLB Act, the State of Texas deems its inclusion in the state tests important. Students in the 8th, 10th, and 11th grades are required to take the social studies TAKS test. However, the 8th grade and 11th grade tests seem more important as the test in eighth grade is the first social studies test the students take, and the eleventh grade test is a part of the graduation requirements.

Like their peers, LEP students, whether in the ESL program or not, are required to take the TAKS test after a specific period of time. Unfortunately, LEP students in the State of Texas have been underperforming in the TAKS tests in general, and in the social studies TAKS test in particular, as shown in the LEP passing rates available in the Academic Excellence Indicator System (AEIS) reports. In the 2002-2003 school year, the overall passing rate for LEP students for 8th graders was 24.3% and for 11th graders 15.2% (TEA, 2003a). In 2006, LEP students performed poorly as well; in 8th grade, the passing rate was 12%, and in 11th grade it was 16% (TEA, 2006a). Finally, in the 2008-2009 school year, the passing rate was somewhat higher with 24% for 8th graders, and 26% for 11th graders (TEA, 2009a). As for the social studies TAKS test, more LEP students met the standards. In 2003, the passing rate was 68.5% and 60.9% for 8th graders and 11th graders respectively (TEA, 2003a). As for 2006, 46% of 8th graders met the standards, whereas for 11th graders the rate was 65%. Finally, an increase of passing

rates was evident in the 2008-2009 school year, with 69% for 8th graders and 79% for 11th graders.

While research to assist ESL students in the classroom to increase performance on standardized tests such as the TAKS has been conducted, there seems to be a gap in the literature in identifying specific areas in social studies TAKS that ESL students find most difficult and challenging. Four of the five test objectives in social studies TAKS embody four major areas of social studies in general: history, geography, politics, and economic and social influences on historical issues and events. The fifth test objective focuses on social studies skills. Since a specific number of test items is allotted for each of these objectives, it would be of value to the teacher to know which areas their students are finding more challenging.

Furthermore, there is a gap in the study of social studies in relation to question test item types, especially those with seemingly more text than the other questions. In examining the question items on the social studies TAKS, success in answering two types of questions appears to require greater textual understanding than other questions. These questions incorporate text boxes with quotes, ideas, and facts. Since linguistic competence in academic English is one aspect of the language that needs to be attained by ESL students in order to succeed in such standardized tests, more language-intense texts in question items must be examined. The nature and organization of these question items need to be explored further to better assist social studies teachers help their ESL students to perform better on social studies standardized tests.

Purpose Statement

The purpose of the study is to analyze the performance of ESL students who took the 8th and 11th grades social studies TAKS test in relation to content and item format to determine if there is a relationship between ESL skills and the students' ability to successfully pass the social studies TAKS exam. The content is reflected in the objectives of the test: history, geography, economics and social influences, political influences, and social studies skills. This study explored the areas in the social studies TAKS that are more challenging than others to ESL students. Moreover, it examined the performance of ESL students in relation to item format; specifically in questions that include reading texts whether excerpt-based or in the format of bullet points.

To further confirm the results of this study, an analysis of longitudinal performance was conducted by examining the scores, in specific categories, of ESL students who took the test in 8th grade and were promoted to 11th grade after three years while still participating in an ESL program. The test scores of students who were in 8th grade in the 2002-2003 school year and were promoted to 11th grade in the 2005-2006 school year, and those of students who were in 8th grade in the 2005-2006 school year and were promoted to 11th grade in the 2008-2009 school year, were examined in order to determine if a pattern exists in performance as a function of maturation of the students.

Finally, demographic effects were explored in terms of gender, ethnic classification, and social economic status (SES) represented in this study as

economically disadvantaged or not, and their effect on student performance in social studies TAKS overall.

Research Questions

1. Is there a significant difference in the performance of ESL students across the five objectives of the social studies TAKS exam?
2. Is there a systematic difference in the performance of ESL students who remained in the ESL program across objectives in the social studies TAKS exam? (Meaning, does the performance of ESL students, promoted from 8th grade to 11th grade, across objectives change over the years?)
3. Is there a significant difference in the performance of ESL students across item formats which include reading texts (excerpt-based, and bullet-point format) in the social studies TAKS?
4. Is there a systematic difference in the performance of ESL students who remained in the ESL program on test items that are excerpt-based? (Meaning, does the performance of ESL students, promoted from 8th grade up to 11th grade, on excerpt-based questions change over the years?)
5. Is there a significant difference in the performance of ESL students in the social studies TAKS exam when specific demographic variables are used as independent variables?
 - a) Does gender affect the performance of ESL students on the social studies TAKS?
 - b) Does ethnic classification affect the performance of ESL students on the social studies TAKS?
 - c) Does social economic status (SES) affect the performance of ESL students on the social studies TAKS?

Limitations

The first limitation to this study lies in the fact that it used secondary data analysis, limiting the data to what is available and what was provided to the researcher. The data available limited the study by the following items. First, only scores of ESL

students in the State of Texas who completed the social studies TAKS test were included in the study. Second, the study is limited to the information provided by the Texas Education Agency (TEA) that did not include detailed information about the students, which might have affected the scores on the social studies TAKS tests. Specific student information such as age of arrival in the United States, length of residence in the United States, and degree of education in their homeland or the amount of formal education in their native language, have been cited as factors in understanding student performance (Cho & Reich, 2008; Collier & Thomas, 1989; Curtin, 2005). These factors that may affect both the speed and success of the students' second language acquisition were not available to the researcher.

In addition, this study is limited by a lack of measure of English language proficiency levels of the students, the amount of prior content-area knowledge these students possess, their families' educational background, and finally the "students' former exposure to Western/urban lifestyles" (Cho & Reich, 2008, p. 236).

This study is also limited to the information provided for the years which have corresponding released tests. TEA does not provide released TAKS tests for every year; thus, the study is limited to the information of students in those years of which released tests were provided to facilitate question item identification. Consequently, this study did not include the most recent data (2009-2010 and 2010-2011 school years) on students and their performance as they had not been released.

Delimits

The researcher chose to analyze information on ESL students only in 8th grade and 11th grade, excluding 10th grade students who took the social studies TAKS test. As mentioned earlier, the reason being is that the 8th and 11th grade social studies TAKS results have consequences to the students. The 8th grade social studies TAKS is the first social studies state test to be given to students, and although its scores do not yield any major decision such as promotion to the next grade, it is a milestone and an indicator of a student's performance in the content area of social studies. As for the 11th grade, passing the social studies TAKS is a requirement for graduation.

In addition, the study is limited to the scores of students who took the state TAKS exam, excluding special education students who took the Alternative TAKS exams and any sort of accommodated TAKS exams. In addition, the LEP status was disregarded as ESL students were considered a group to be explored regardless of the level of English proficiency in which each student was categorized. A student may still be considered an LEP student even after he/she leaves the ESL program (TEA, 2012a); therefore, it is not the purpose of this study to track students who exited the program and those who remained even if they were labeled LEP, rather to analyze the data of students who participated in an ESL program. Thus, the researcher disregarded the information provided regarding the LEP status of students.

Finally, in terms of question item analysis, this study is limited to two types of text-based test items: excerpt-based questions and bullet-point questions. After examining sample questions of the social studies released TAKS tests, two types of text-

based questions that are enboxed and organized on the page differently, were identified. This study aimed at exploring the performance of ESL students on test items with relatively large amounts of text, specifically in terms of the organization of the text. Therefore, other question items were not included in the study although they may have included more text; that is because the text in these question items is part of the question sentence rather than provided as additional text in which the student had to base the answer on.

Definitions

Academic Excellence Indicator System (AEIS) is a system of information on the performance of each school and district in Texas that is collected in an annual report. The report includes information such as results of the TAKS, attendance rates, and participation in the TAKS tests. Performance on each indicator is shown by ethnicity, gender, special education, low income status, limited English proficiency status.

Bullet-point items: Items containing a box with written text organized with bullet points, each including at least one piece of information. The item is introduced by the following phrase in the test booklet, “use the information and your knowledge of social studies to answer the following question.”

Economically disadvantaged students refer to students who are eligible for free or reduced-price meals under the National School Lunch Program and Child Nutrition Program or other public assistance (TEA, 2009b). The code indicates

that the student's family has an income low enough for the student to qualify for one of these programs.

English Language Learner (ELL) is an alternative to the term LEP, as it refers to students whose first language is not English, regardless of their proficiency level (beginner or intermediate levels).

ESL (English as a Second Language) students: Refers to students whose primary language(s) is other than English. Therefore, they require additional services to help them develop their English language proficiency; thus, they are enrolled in an ESL program.

Ethnicity is a term used to classify students according to how they identify themselves as being part of a race or a cultural group. The five categories of ethnicity students may choose from for self-identification are Native American, Asian, African American, Hispanic, and White (TEA, 2009b).

Excerpt-based items are items containing a box with written text, a portion excerpted from a longer document that includes an indication to the source of information following the text. The item is introduced by the following phrase provided in the test booklet, "use the excerpt and your knowledge of social studies to answer the following question."

Limited English Proficiency (LEP) refers to students who have not become fully fluent in the English language and are unable to use English effectively to complete grade-level work in a mainstream classroom successfully without support.

NCLB (No Child Left Behind Act) refers to an act of Congress (2001), which establishes high standards and measurable goals to improve education in American schools (U.S. Department of Education, 2002).

Secondary data analysis refers to analysis of an existing dataset that had been gathered (Hakim, 1983), such as data produced by the national Census, or data from tests of student performance...etc.

SES (Socio-economic Status) refers to the social and economic position of the student's family, including household income, education, and occupation.

Social studies achievement is measured by the proportion of correct student responses to items on the social studies TAKS test.

TAKS- Alternative (TAKS-Alt) is an alternate assessment which is designed for students with cognitive disabilities, and who are receiving special education services. The assessment involves observing students as they complete tasks that are linked to the curriculum, and then are assessed using a special TAKS-Alt rubric.

Texas Assessment of Knowledge and Skills (TAKS): A standardized state mandated test used in Texas, administered yearly to assess students in the areas of reading, writing, mathematics, science, and social studies under Texas education standards.

Texas Education Agency (TEA) is an agency comprised of the commissioner of education and other staff members. The TEA and the State Board of Education guide and overlook the educational programs and activities of public education in

the State of Texas, with the goal of helping schools meet the educational needs of all students.

Texas Essential Knowledge and Skills (TEKS) refer to a set of standards that describe what students in Texas schools should know and be able to do. They were adopted as state curriculum effective September 1998.

Title I Schools are elementary and secondary schools that are provided with financial assistance because of the enrollment of high numbers of low-income students in them, with the aim of ensuring that all children meet the state academic standards (U.S. Department of Education, n.d.).

Organization of Chapters

The current study analyzed ESL students' performance on the social studies TAKS test in grades 8 and 11, for the years 2003, 2006, and 2009. This dissertation is organized in five chapters. Chapter I, the introduction, includes a statement of the problem, the purpose of the study, the research questions that guide the study, limitations and delimitations of the study, and definitions. Chapter II is a review of literature in the areas of accountability and standardized testing with a focus on the TAKS test, social studies and TAKS, content and skills in the social studies TAKS test, the intersection of ESL students and the social studies section of the TAKS test including the challenges ESL students face in this portion, enumeration cues in social studies TAKS test items, and finally a review of literature on the effect of gender, ethnicity, and socio-economic status on student performance. Chapter III, entitled "Design of the Study," presents a theoretical framework and an analytical framework, ending with a description of the

analytic process. Chapter IV includes a detailed presentation and analysis of the data organized in five sections that correspond with the five major research questions, followed by a summary of findings. Finally, Chapter V offers an overall summary of the purpose of the study, a discussion of the results, general implications for both theory and practice, and ends with a discussion of limitations and future recommendations.

CHAPTER II

LITERATURE REVIEW

In the United States, the number of English Language Learners (ELLs) has increased dramatically in the decade between 1999 and 2009. Approximately, an increase of 2 million ELLs in a ten-year span occurred (National Clearinghouse for English Language Acquisition, 2011). As these numbers increase, the public's interest and policymakers' attention are directed towards them, and how they perform academically especially after the passing of the *No Child Left Behind Act* (NCLB) of 2001 (Flores, Batalova, & Fix, 2012). Per this Act, states are required to assess ELLs both in English language proficiency and in content-area knowledge; subsequently, states are being held accountable for that (Flores et al., 2012).

The State of Texas has the distinction of possessing the second-largest number of ELL students in the nation (California leading with over a million ELLs), with 838,494 ELLs in 2011 (TEA, 2012c). The ELLs group accounts for about seventeen percent of all students (PreK-12) in Texas. This population, like other population groups, went through the process of being incorporated into the U.S. school and accountability systems under the NCLB act. Since accountability is a major part of the NCLB act, ELLs were offered linguistic assistance to enhance their academic performance. ELL students are eligible for English language instruction, in which its form is determined by the state. The process of creating these special language acquisition programs for ELLs had not been simple, as there have been legal and political issues throughout the years in

that regard (Flores et al., 2012). In Texas, these special language programs take the form of either bilingual programs or ESL programs.

To situate the current study in context, a general review of previous studies conducted in the field follows. The review of literature starts with a general view of the concept of accountability, and more specifically in regards to standardized testing. In connection to that, the TAKS test is discussed extensively. Within the discussion of the TAKS test, special focus is placed on the social studies section of it. Since the objectives of the social studies TAKS test are divided into two major areas of content and skills, these are explored as well. Later, the intersection of ESL students and the social studies section of the TAKS is explored while emphasizing the challenges that ESL students encounter with the social studies part of the TAKS test, both linguistically and in content. Additionally, social studies TAKS test items that include enumeration cues are discussed. Finally, a review of literature concerning the influence of demographics on academic achievement is presented; the effect of gender, ethnicity, and socio-economic status on student performance is analyzed.

Accountability and Standardized Testing

With the passage of the No Child Left Behind Act (NCLB) in 2001, testing has gained its status in the spotlight in terms of accountability. Standardized tests in particular are used to ensure the ‘high standards’ education of all students (Coltrane, 2002; Eslami & Plett, 2008). These tests are considered high stakes as the scores they generate are used to determine major decisions in a student’s education, whether it be about placement, promotion, or graduation (Coltrane 2002; Heubert and Hauser, 1999).

In the State of Texas, to comply with the NCLB Act, the standardized Texas Assessment of Knowledge and Skills (TAKS) test was administered yearly to all the students. Every student had to take the TAKS test in order to be promoted to the next grade level (in grades 3, 5, and 8), or to graduate (exit level/11th grade) (Burroughs, Groce, & Webeck., 2005; Pyron, 2007). Recently, however, the TAKS test was replaced by the State of Texas Assessments of Academic Readiness (STAAR) beginning in the Spring 2012 semester (TEA, 2012b). The STAAR program differs from the TAKS as it focuses on increasing “the rigor of both the assessments and the performance standards for all grades, subjects, and courses,” along with other differences (TEA, 2012b, paragraph 4). However, the change is most prominent at the high school level as the content areas tested are “replaced with twelve STAAR end-of-course (EOC) assessments” (TEA, 2012b, paragraph 1), whereas the same grades and subjects for grades 3-8 as the TAKS are still tested. Social studies will still be included in the STAAR testing both in the eighth grade and in the high school level (with a different format). At the high school level, the twelve EOC assessments include World History, World Geography, and U.S. History (TEA, 2012b).

In particular, in terms of its accountability system, curriculum, and assessment, Texas has been in the lead compared to other states (Ikegulu, 2004). In fact, as accountability is defined to be the system in which schools are monitored in terms of their teaching and their students’ performance according to the state-required curriculum, Texas was “the model for the federal accountability system established in the *No Child Left Behind (NCLB) Act* of 2001” (Ikegulu, 2004, p. 4). That holds true as

observers concluded in the late 1990s that the accountability system in the State of Texas, which was test-based, had positive results in terms of educational progress (Haney, 2001).

States were required to administer tests to students in grades 3 through 8, as a result of the NCLB act (Chargois, 2008). Each state decided which grade levels and which tests were to be passed in order to be promoted or to graduate. In Texas, the passing of tests in the third, fifth, and eighth grades determined if the students were to be promoted, and in the eleventh grade, passing the TAKS tests was a requirement for graduation (TEA, 2009c). The TAKS test is basically the only thing “the state controls totally” (Pyron, 2007, p. 17). According to Pyron (2007), students may be able to get through without meeting all the other requirements of graduation, but not the TAKS requirement. It is worthy to note here that the Commissioner of Education in the State of Texas allowed an alternate assessment for students who transferred from other states after the middle of the year in which they were expected to graduate. The Scholastic Assessment Test (SAT) and the American College Testing (ACT) verbal and mathematics tests may have substituted for their TAKS sections, but students still had to pass the exit level science and social studies TAKS for them to meet the requirements of graduation (Dietz, 2010).

Whether promotion or graduation, these are major milestones in a student’s educational career. As these milestones are reached through tests such as the TAKS, these tests came to be called high stakes tests (Phelps, 2003). No matter how well a student performed in school, or whether he/she met all the other requirements for

graduation, the gatekeeper for graduation in the State of Texas was the TAKS (Pyron, 2007). High stakes tests are tests that have punitive effects on the students and the schools in which these students study (Cizek & Burg, 2006; Ikegulu, 2004; U. S. Department of Education, 2002). In fact, the performance of the students on these tests determines the success of the school, the district, and even the state (Bouroughs et al., 2005; Chargois, 2008; Coltrane, 2002; Funkhouser, 1990). Because of that, student performance on standardized tests is being linked to different variables such as teacher or administrator capability, school or class size, ethnicity, district adequacy, and other variables that may affect the performance of the students (Black & Valenzuela, 2004; Chargois, 2008; Koretz, 2002).

The Texas Assessment of Knowledge and Skills (TAKS)

The TAKS test was first administered in the 2002-2003 school year, after approximately three years of development (Cole, 2007). The test was designed to test the Texas Essential Knowledge and Skills, known as the TEKS, which basically describe the curriculum. Both the TEKS and the TAKS are included under the Academic Excellence Indicator System (AEIS) in Texas (Burroughs et al., 2005). A list of TEKS exists for every grade level in every subject taught; however, not all the TEKS are covered in the TAKS test; rather, the ones chosen every year change (Pyron, 2007). Along with English Language Arts (ELA), mathematics, and science tested at different grades, students are given the social studies TAKS in grades 8, 10, and 11 (exit level) (TEA, 2007a). Social studies TAKS, like the other subjects tested, is divided into

objectives. These objectives, which are the same for all grade levels, are general statements that underlie the student expectations derived from the TEKS (TEA, 2004a).

The first statewide assessment in the State of Texas dates back to 1980, but the accountability system, the Academic Excellence Indicator System (AEIS) which emphasizes student achievement, was created in 1984 (Ikegulu, 2004). To meet the demands of the accountability system, states were required to develop more rigorous mandated standardized tests (U. S. Department of Education, 2002). In the State of Texas, tests are developed from within to ensure the curriculum requirements are met (Ikegulu, 2004). The state-mandated curriculum is the TEKS, which are basically curricula standards that guide classroom teachers through their instruction (TEA, 2004a). Accordingly, the curricula implemented in the public schools of Texas are to be based on the TEKS. Therefore, a decent mastery of the TEKS is needed in order to perform well on state-mandated tests (Ikegulu, 2004).

From 1990 until 2002, the Texas Assessment of Academic Skills (TAAS) served as the state assessment. However, on June 18, 2001, the TAKS test replaced the TAAS test (Pyron, 2007). According to the 8th grade Information Booklet (TEA, 2004a), the TAKS program is a completely renovated testing program as it includes more TEKS than its predecessor and uses more authentic questions. The TAKS test was developed with the purpose of:

1. meeting a state mandate
2. providing schools with student academic diagnostic information
3. determining prospective high school graduates' mastery of the state curriculum
4. encouraging districts and schools to identify and serve students at risk of academic failure

5. promoting equity of opportunity across all student groups, including English language learners (ELLs)
6. determining prospective high school graduates' knowledge and skill levels in English and mathematics relative to those needed for postsecondary education
7. providing data to state policymakers on student attainment of state education goals to inform education policy decisions.

(Dietz, 2010, paragraph 2).

The TAKS test is a criterion-referenced standards-based test, in which the attained score by the test takers is compared to a curriculum (in this case, the TEKS) (Ikegulu, 2004). As mentioned earlier, the test questions on the TAKS test are closely aligned with the TEKS; therefore, the development of the test included various stake holders in order to ensure the adequacy and accuracy of the measurements (TEA, 2004a). Teachers, parents, educational professionals, both Texan and national content-area experts including college and university faculty were involved in developing the TAKS program within a time frame of three years beginning in 1999 (TEA, 2004a). The first step was to identify the TEKS for each grade and for each subject area that was to be assessed. Surveys were created to review the TEKS and the objectives for each subject area, and were then sent to Texas educators, yielding more than 27,000 survey responses; the TAKS was adjusted based on the input received from the surveys (TEA, 2004a). In the center of the TAKS lies the concept of vertical alignment to ensure that the TAKS test was going to be more rigorous as a student progressed through grades (TEA, 2004a). That is, to see that academic skills are being developed in alignment with the TEKS, consequently ensuring a strong enough foundation for success (TEA, 2007a).

The TAKS is a testing program administered to public school students every year, usually in the spring, starting in the third grade and ending in the eleventh grade

(TEA, 2007a). Six major subjects are tested using the TAKS: mathematics (grades 3-11), reading (grades 3-9), writing (grades 4 and 7), English Language Arts (grades 10 and 11), science (grades 5, 8, 10, and 11), and finally social studies (grades 8, 10, and 11) (TEA, 2007a). The respective tests measure how much the student has learned and is able to apply what is expected of him/her in terms of concepts and skills (TEA, 2007a).

The TAKS test was first administered in spring of 2003; however, only in 2005 did the State begin withholding diplomas of students who did not pass the TAKS tests (Dietz, 2010). To graduate, students are required to pass the four content areas: mathematics, ELA, science, and social studies. Although the TAKS test was not meant to become the ‘gate keeper’ for graduation, it ended up being so (Pyron, 2007). However, students in eleventh grade who have met all the other graduation requirements, but have failed the TAKS, may retake the exit-level tests up to five times before they complete their senior year; after that, there is no age or number limit on retakes (Dietz, 2010).

Social Studies and TAKS

Although social studies is not included as a subject that is required to be tested according to the NCLB act, the State of Texas includes it in the TAKS tests (Burroughs et al., 2005). It is worth noting, however, that not all states include social studies in their state exams, and there are conflicting feelings and concerns in that regards from teachers in the field (Burroughs et al., 2005). Some states, such as Illinois, even eliminated social studies state tests after the passage of the NCLB act (Manzo, 2005). Nonetheless, the State of Texas believes in the importance of this subject matter as its assessment

program establishes the social studies test as a requirement for graduation (Pyron, 2007; TEA, 2004a).

Flanagan and Faison (2001) point out the importance of the adolescent period in one's life in developing a civic identity as a student takes on more responsibilities as an active citizen. Regardless of the fact that social studies is a content area that is being assessed in the TAKS tests, not much attention has been given to it in comparison with other subjects such as mathematics or ELA. According to Westheimer and Kahne (2003), the majority of school resources are being directed towards subjects such as mathematics, science, and literacy. Moreover, Bourroughs et al. (2005) explain that because social studies is tested in the secondary level, elementary level students are receiving less instructional time in social studies, and teaching resources are decreasing in favor of the other subjects. Bedford (2007) echoes that by saying that insufficient funds are allocated for civic education programs in public schools.

Not only are social studies knowledge and skills a graduation requirement, the study of social studies has its importance in developing an educated citizen of his/her country (Burroughs et al., 2005). It is a study that "helps students understand their place in the world and their role as responsible citizens" (TEA, 2004a, p. 4). Additionally, a number of skills are acquired and developed through the study of social studies including critical thinking skills (TEA, 2004a). The importance of social studies is also mentioned in the Texas Education Code, Chapter 4, (2009) where it is stated that educators are to prepare students to be active citizens who appreciate the values of both the state and the national heritage, and finally to teach them how to function in the society. Nevertheless,

because social studies is a young subject in high-stakes standardized testing, the impact of that entry in the high-stakes assessment system is yet to be explored, as it is “only emerging” (Burroughs et al., 2005, p.16).

The questions on the social studies TAKS are in the multiple-choice format, and the test is divided into test objectives (TEA, 2007a). Social studies objectives are derived from the TEKS, as with objectives in other content areas tested on the TAKS (Bedford, 2007). The objective statements group student expectations, derived from the TEKS, in order to organize the test into reporting units that are easily analyzed by stake holders such as campuses and parents (TEA, 2007a). The analysis of these reporting units helps different entities understand the performance of students and their schools (TEA, 2007a). These objectives are broad statements that do not change across grade levels. Thus, the social studies objectives are identical for eighth, tenth, and eleventh grade (TEA, 2007a). However, the topics are developmentally appropriate as some topics tested in grade 8 are repeated again in grades 10 and 11 (TEA, 2007a).

The social studies TAKS test is divided into five objectives:

1. understanding of issues and events in U.S. history.
2. understanding of geographic influences on historical issues and events.
3. understanding of economic and social influences on historical issues and events.
4. understanding of political influences on historical issues and events.
5. using critical-thinking skills to analyze social studies information.

(TEA, 2004a).

These objectives are valid for all three levels, eighth, tenth, and exit level in social studies. Study guides for the social studies TAKS are usually organized following the sequence of the five objectives. Table 1 explains the five objectives, what is expected of the student for each, and examples of topics that may be assessed in each objective.

Table 1*Social Studies Objectives in the TAKS Test, Student Expectations, and Sample Topics.*

	Student Expectations	Sample Topics
Objective 1	Demonstrate an understanding of issues and events in U.S. history	Important dates in early American history and the roles that George Washington and Thomas Jefferson played in the American Revolution
Objective 2	Demonstrate an understanding of geographic influences on historical issues and events	How maps, charts, and graphs communicate geographic information and how geographic factors influenced major events in U.S. history
Objective 3	Demonstrate an understanding of economic and social influences on historical issues and events	The development of the free-enterprise system and the importance of technological innovations in U.S. history
Objective 4	Demonstrate an understanding of political influences on historical issues and events	The development of representative government in colonial America and fundamental principles of the U.S. Constitution and the Bill of Rights
Objective 5	To use critical-thinking skills to analyze social studies information	How to interpret written and visual sources of social studies information

(TEA, 2007b)

The TAKS tests are organized within blueprints that determine details such as the number of test items for each objective, in order to ensure the representation and distribution of the TEKS across objectives (TEA, 2004a). Table 2 summarizes the

objectives and the number of questions related to each objective (which did not change in 2003, 2004, 2006, and 2009).

Table 2
The Number of Questions for Each Social Studies Objective.

TAKS Objectives	Grade 8	Grade 10	Grade 11
Objective 1: History	13	7	13
Objective 2: Geography	6	12	9
Objective 3: Economics	9	7	13
Objective 4: Political Influences	12	12	9
Objective 5: Social Studies Skills	8	12	11
Total number of items	48	50	55

(TEA, 2004b)

The social studies TAKS test objectives serve as a grouping tool to ease the organization of the test. However, according to Bedford’s (2007) study, the five objectives are closely related in terms of student performance. The study of 1,319 eighth grade students who passed the TAKS test in 2005 at two school districts in South Texas, suggested that those who scored high on civics knowledge, represented by Objective 4, also scored high on history (Objective 1) ($r=0.57, p<0.001$), for example. In fact, Bedford (2007) found that civic knowledge (Objective 4) scores were positively associated with the other four objectives’ (history, geography, economics, and social studies skills) scores and the associations were statistically significant at the 0.001 level ($r=0.57, r=0.43, r=0.44, r=0.42$ respectively). Moreover, the TAKS Information Booklet for the 8th grade social studies test states that the 8th grade social studies test

requires a strong base in social studies skills and content knowledge that have been accumulated in the elementary years.

Content Knowledge versus Skill

As mentioned previously, the five objectives of the social studies TAKS test include history, geography, economics, political influences, and social studies skills (TEA, 2007b). The five objectives may be categorized as either content/knowledge objectives or skills objectives (Smith, 2012). The fifth objective maintains an emphasis upon student skills, as the name implies (Smith, 2012), rather than emphasizing student knowledge as is demonstrated in this section.

Elementary, middle, and high school social studies knowledge and skills TEKS are divided into different strands including history, geography, economics, government, citizenship, culture, science, technology, society, and social studies skills (TEA, 2011b). These strands include objectives that start with the phrase “the student understands....,” and is then followed by the objective, which is followed by specified tasks the students are to master (Smith, 2012). These tasks begin with one of the action verbs in Bloom’s Taxonomy list (define, describe, identify, explain, summarize, trace, apply, analyze, compare, contrast, and evaluate) (Smith, 2012).

Smith (2012) analyzed the social studies curriculum extensively by examining the objectives in the K-12 social studies standards. He specifically focused on the 2010 Texas state social studies standards. Through his analysis, he found that although the curriculum is intended to focus on critical thinking skills, in reality much of the focus is on memorization. The lower verbs of the Bloom’s Taxonomy were used heavily; in the

middle school level, the majority of the verbs used were below the level of critical thinking (Smith, 2012).

The analysis of the eighth grade social studies strands revealed that the first four major objectives of the social studies TAKS test focus on knowledge in general (Smith, 2012). In history, the focus is on specific historical events, eras, locations and names...etc. Factual knowledge is emphasized, which, according to Smith (2012), could be easily assessed by standardized tests. This was the case with all the other strands for both the middle school and the high school levels.

However, the last strand (social studies skills) for the TEKS in middle school includes objectives that incorporate more verbs from the upper levels of Bloom's Taxonomy; therefore, shifting the emphasis towards developing student skills (Smith, 2012). In this strand, a student is expected to "create thematic maps, graphs, charts, models, and databases representing various aspects of the United States" (Smith, 2012, p. 184); rather than being asked to learn and memorize information, the student is prompted to create and produce. The strand's encompassing nature helps students master their skills via cross curricular knowledge. The incorporation of skills from other content areas such as mathematics, science, and language arts reinforce student skills developed in the classroom (Smith, 2012). An example of cross curricular knowledge incorporates mathematics skills by expecting students to "use appropriate reading and mathematical skills to interpret social studies information such as maps and graphs" (Smith, 2012, p. 358). Overall, in this strand, the objectives begin with phrases such as "The student communicates," "The student applies," or "The student uses," followed by "how the

student must communicate and what the student must use to achieve the standard” (Smith, 2012, p. 284).

Whether the objectives are of the first seven strands or the eighth strand may influence the way ESL students perform due to the different nature of these strands. An emphasis on content knowledge versus skill could have an influence on how well ESL students may answer test questions in the social studies TAKS.

ESL Students and Social Studies TAKS

A student of limited English proficiency (LEP) is a student whose English language skills limit his/her performance in ordinary class work, and whose primary language is other than English (Texas Education Code, n.d.). An ESL student, however, is a student who participates in LEP programs in grades 9 through 12. ESL students receive instruction in the purpose of developing their English language skills, focusing on the TEKS for reading and language arts (TEA, 2012a). The placement of students in the ESL program starts with the student indicating that he/she speaks a language other than English at home in the home language survey, then the state requires an oral test to be administered, and based on the results, the student either qualifies or not for the ESL program at the secondary school (Lewis-Moreno, n.d.). However, in the elementary school, students are placed in the bilingual program.

ESL students are tested regularly; they take tests such as the Scholastic Assessment Test (SAT) and the TAKS in addition to a number of tests that assess their academic development and English language acquisition (Pyron, 2007). An ESL student has to pass the reading and writing parts of the TAKS and show oral proficiency in

English in order to exit the ESL program. He/she is then monitored to check for additional support if needed. In relation to TAKS, generally students qualify for one year of exemption in their first year in the U.S., and may qualify for as long as three years depending on a committee's decision; nevertheless, high school students cannot be exempt from the exit level TAKS tests, but may postpone them (Lewis-Moreno, n.d.). Overall, the duration the student remains in the ESL program is an educational and political matter that is determined by the state and varies from one state to another, or may vary from one year to the next (Duff, 2001). This limitation of the exempt status of LEP students places burden on both the students and their teachers (Curtin, 2005).

ESL students in the secondary level are vulnerable especially due to the fact that second language acquisition is more challenging and difficult as students get older (Cho & Reich, 2008). Furthermore, as mentioned earlier, the number of tests ESL students have to take places more stress on them and can create frustration resulting in the feeling of yet more isolation from the other students (Cho & Reich, 2008). These standardized tests generate yet another challenge of meeting the higher levels of academic achievement that are expected of them, as these tests aim at raising standards for student learning (Coltrane, 2002). Furthermore, scarcity in time is another issue secondary level students encounter. Students in the secondary level have a short period remaining in school before they graduate, which makes it more challenging for them to master basic English language skills while receiving content-area instruction (Collier & Thomas, 1989).

As mentioned earlier, the study of social studies is essential in creating responsible citizens who can be productive members of the society. However, this is not something that is taught in a class or a course throughout the year, but rather a process that develops gradually starting with kindergarten education (Duff, 2001; Smith, 2012; TEA, 2004a). ESL students would have missed out on all the knowledge other students have been exposed to in earlier years, putting them at a disadvantage (Cho & Reich, 2008; Duff, 2001). Both the skills and content of social studies are developed throughout the years and are built upon the materials learned in previous grades (TEA, 2007a); thus, creating a challenge for those who enter the educational system at later phases of their education.

In the State of Texas, ESL students are required to take the TAKS test three years after their arrival in the United States based on the assumption that this time period is sufficient for them to pass standardized tests (Curtin, 2005). Overall, not until recently is more attention being given to the needs of ESL students in the social studies classroom (Case, 2006). However, studies of second language acquisition, such as those of Cummins (1979), demonstrate the inadequacy of this amount of time given to ESL students before taking the TAKS tests.

Cummins (1979) distinguishes between what he calls Cognitive Academic Language Proficiency (CALP), and Basic Interpersonal Communication Skills (BICS). The first, CALP, which takes longer to achieve, is associated with higher-order thinking skills such as analysis and synthesis, normally essential for success in a social studies class. Five to seven years are required to gain CALP in a second language (Cummins,

1981), or even five to ten years according to Collier and Thomas (1989). The number of years it takes a student to acquire cognitive-academic second language (L2) proficiency varies by person and may depend of factors such as subject mastery in the first language (L1) (Collier & Thomas, 1989). In sum, “[c]ognitive-academic second language proficiency,” as Collier and Thomas call it, is a gradual process that necessitates a number of years (Collier & Thomas, 1989, p.34).

Thus, after three years of ESL instruction, ESL students are most likely not ready for a standardized test in English as they have not developed the second language proficiency needed for academic purposes (Collier & Thomas, 1989). However, after the state specified amount of time (three years) for ESL students to be ready for the high-stakes standardized test, TAKS, most students will have only achieved the BICS, which refers to basic language skills including simple asking and answering of questions, and conversing about everyday life (Case, 2006). Curtin (2005) attests to that in her study of six middle school ESL students in their third year in the United States; the six students were not successful in passing the standardized tests they took that year, according to the study. The problem, as Pyron (2007) states, is that ESL students can pass their classes at school with the help of their teachers and the accommodations offered by their schools, but they are unable to pass the TAKS when this support is not available.

Another issue raised by Menken (2000) is the lowered validity of the tests for ESL students; meaning, ESL students are not assessed in terms of their content knowledge, but rather their English language proficiency. Coltrane (2002) echoes that by saying that if students are unable to demonstrate their knowledge of the content areas on

tests due to linguistic difficulties, then the validity of the results are to be questioned; the low scores of students on standardized tests may simply be the result of not mastering the English language skills to take the test and not being able to utilize the language to demonstrate content-area knowledge (Case, 2006; Coltrane, 2002). This may specifically be true in the content area of social studies, as it requires a high level of literacy (Gonzales, 1988).

Within the TAKS test, yet another challenge, related to linguistic competence, arises: text type. Indeed, both item format along with test content affect test performance (Bachman & Palmer, 1982). Several research studies have been conducted to explore the impact of text types or genres on students' performance, especially in the area of reading passages and language arts tests (e.g., Carrell, 1984; Kobayashi, 2002). Second language learners must develop lexical knowledge before they read in order to process texts that encounter them (Shin, 2002). Several factors influence the 'reader-text interaction', as Shin (2002) calls it, including the reader him/herself, and the text. The text factor incorporates several aspects such as the language of the text, which is the major aspect/variable (Bernhardt & Kamil, 1995; Shin, 2002), the content of the text, and text organization (Shin, 2002). In general, studies that explored the effect of text structure on readers' performance found that a reader's prior/background knowledge played a major role (Eslami & Plett, 2008; Shin, 2002). The role of this knowledge is known as schema theory (Carrel, 1984). This theory holds that readers use a prior schema they possess in order to interpret a text. This schema may be affected by culture (Carrel, 1984),

organization of the text (Carrel, 1985), topical structure, and content domain of the text (Carrel, 1987).

Some challenges ESL students may face include the complex social studies vocabulary (Chamot & O'Malley, 1994; Cho & Reich, 2008; Gonzales, 1988), and sentence structure (Cho & Reich, 2008). Other challenges are related to cultural differences in terms of understanding and analyzing concepts (Cho & Reich, 2008; Coltrane, 2002; Duff, 2001; Eslami & Plett, 2008; Reilly, 1988) and lack of background knowledge in the content area (Chamot & O'Malley, 1994; Cho & Reich, 2008; Duff, 2001). Finally, language barriers and the need to attain a high level of literacy to understand the curriculum (Coltrane, 2002; Duff, 2001; Eslami & Plett, 2008; Gonzales, 1988) are challenges that ESL students may face in the content area of social studies.

The social studies curriculum, specifically, includes a large amount of content-specific vocabulary, namely social studies concepts (Cho & Reich, 2008). The challenge of this content-specific vocabulary lies in that it assumes rather extensive background knowledge, and may require more advanced language skills in terms of reading and writing (Cho & Reich, 2008). The increasing difficulty in vocabulary with the advancement of grade levels is a fact that ESL students have to deal with. The complexity of the concepts and what they represent increases with grade level, thus, increasing the difficulty (Gonzales, 1988). In test items, these terms and their references or ideas may not be familiar to the ESL students simply because of their lack of exposure to such terms, or their possession of a different understanding of the terms (Coltrane, 2002). According to Chamot and O'Malley (1994), the highly abstract nature

of the social studies concepts makes them more difficult for ESL students to recognize. Gonzales (1988) adds that the social studies vocabulary is developed from abstract ideas that are part of areas such as philosophy, anthropology, and economics among others, which increases their difficulty. Moreover, this discipline-specific vocabulary, including words such as *embargo*, is not really part of an average person's everyday life, which makes it difficult for a student to understand their meanings (Cho & Reich, 2008). Moreover, certain vocabulary words in the social studies curriculum may have different meanings than what is used in everyday life; for example, the word *act* holds a whole new perspective to the word when it is mentioned in social studies materials (Cho & Reich, 2008); generally, each concept, *democracy* for example, stands for a set of ideas rather than having a simple explainable meaning (Gonzales, 1988).

Also, related to the vocabulary found in the social studies curriculum is the effect of culture. First, social studies instruction is not necessarily a content area that is taught and learned in every educational system around the world; some students may have not been exposed to any social studies instruction in their countries (King, Fagan, Bratt, & Baer, 1987). Second, one disadvantage ESL students may have is that the content in which they are extensively familiar with such as the history of their country, Chinese history for example, is seldom included in the social studies curriculum in the U.S. (Duff, 2001; Smith, 2012). Third, even if ESL students have had social studies instruction in their native languages, some concepts or ideas, such as the concept of reaching compromises to make decisions, may be entirely new and never heard of to them (Reilly, 1988). Fourth, culture plays a role in understanding the differences in

meanings; the word *government*, for example, may allude to different things and may be understood differently (Cho & Reich, 2008). Interestingly, the lack of background knowledge of content area accounted for 70.6% of teachers' concerns with ELLs in Duff's (2001) study, which was conducted in Canada at a secondary school with a large ESL population from Asian origins.

General knowledge of the popular culture of North America was deemed a necessity for successful and effective participation in a content-area classroom (Duff, 2001). In fact, more recently, the social studies curricula have been focusing on viewing history and historical events from different viewpoints (i.e. critically looking at history), which is an educational method that may not be familiar to students from other cultures (Duff, 2001). This critical view may even be looked down upon by educational systems of other cultures.

To add to the cultural aspect, other than with discipline-specific vocabulary, test-taking skills and formats may be a result of cultural influences. ESL students must be familiarized with the format of standardized tests (both in relation to the language used in the tests, and to the general layout and organization of them) in order to understand the nature of the test to perform better on standardized tests (Coltrane, 2002). ESL students may not have prior experience with such tests, even in terms of test-taking skills especially with multiple-choice items (Coltrane, 2002).

Social Studies TAKS Test Items and Enumeration Cues

Test items on the social studies TAKS include questions that include boxed texts. Of the question types in the test that incorporate boxed texts are those which

incorporate excerpts, and those which utilize enumeration, as shown in the sample questions of study guides and in released tests (TEA, 2007a; TEA, 2007b). Although both question types, excerpt-based and bullet-point questions, share the typographic device of enboxed text, they differ in text organization inside the boxes.

The layout and design of educational materials, including tests, are important as they “affect the way in which students are able to extract meaning from educational texts” (Narveson, 2001, p. 3). According to Diaz (1995), the structure and design of texts affect cognitive performance. Cognitive activities that involve skills such as recalling are associated with how the information is presented, i.e. how the text is designed (Bernhart, 1986; Diaz, 1995; Lorch & Chen, 1986; Narveson, 2001). How learners cognitively process information is affected by signals that are used in the design of a text (Lorch & Chen, 1986). These signals or cues, when used in texts effectively, may facilitate or hinder the learner and his/her performance (Narveson, 2001).

According to Britton, Glynn, Meyer, and Penland (1982), semantic and structural variables determine the meaning of the text. However, the semantic variable is more concerned with the actual meaning of the content. On the other hand, structural variables include the organization of the text which is independent of the content, yet directed to facilitate the meaning of the content (Britton et al., 1982). One example of structural variables is typography. Typography is a form of textual representation (e.g. enumeration) that enhances the content without changing its meaning (Narveson, 2001). Typographical signaling devices include highlighting, italics, color, bulleted points, and enumeration, which emphasize certain aspects of the content without adding any

semantic content. These devices may signal content, such as titles, others may cue organization such as enumeration, and some may signal both (Narveson, 2001).

Text design or presentation technique is a visual (*explicit*) technique that signals the content of the text (*the implicit*), as suggested by Jonassen (as cited in Hayes, 1997, p.3). These visual signals are used by the learners to access the organization and content of the text; thus, being able to comprehend the content of the text (Bernhart, 1986). The complexity of the reading process must be considered in relation to text comprehension. There are many effects that may interact to explain the complex nature of reading (Narveson, 2001). This is one reason research on typography is limited. Each typographic device cannot be examined separately (Diaz, 1995; Narveson, 2001) as it does not function alone; an interaction of these devices and other effects may influence the complex cognitive process of reading. Therefore, cognitive processes such as comprehension and reading are mostly theoretical as they are difficult to measure and prove (Narveson, 2001). There are so many variables such as the different signaling devices, the texts used, and the subjects used and their abilities, which lead to conflicting results of the research in this area (Hayes, 1997). Furthermore, students look at the text in a rather holistic way that includes both content and organization, which further demonstrates the difficulty of singling out one variable in the text, yet favoring the examination of the interaction of those variables (Diaz, 1995).

Of interest to this research are enumeration devices, which include bulleted points and numbered points. Bulleted points use symbols (such as a bullet, a point, an asterisk...) at the beginning of a point. Numbered points, on the other hand, use

numerals *1,2,3...or I, II, III...* or utilize words such as *first, second...*). Enumeration devices indicate organization and may imply an outline; bulleted points, in specific, may suggest significance to the reader (Narveson, 2001). Narveson (2001) also found that order is not necessarily implied by enumeration unless semantic context implies it. In their study of college-age students who were presented with two text versions, one with and one without enumeration, Lorch and Chen (1986) found that enumeration facilitates recall. Recall is increased when the text is enumerated, as enumeration draws the learner's attention to signaled information (Narveson, 2001). Recall, as a measure of learning, is used as it is relatively easy to measure quantitatively (Narveson, 2001).

Also in relation to this study, the effect of typographic signals on readers of different abilities, especially those of lower reading skills is important. That is because these signals may facilitate or, in contrast, hinder the ability of ESL students who are of lower reading skills than their non-ESL peers to answer test questions that include text. The research conducted in this area uses different cues with different measures and different texts, with subjects of varying abilities, which makes it almost impossible to have consistent results (Narveson, 2001). Consequently, the variability of the results in these studies may not be accurately generalized. In Narveson's (2001) study, there were no criteria to specifically classify the subjects into weak or strong readers. However, the findings of the study indicate that typographic cues help weaker readers, but may hinder stronger readers.

As noted earlier, students' learning outcomes in reading, for example, cannot be attributed to one variable such as recall, text organization, or use of typographic devices.

However, these variables may be explored further to complete the bigger picture of incorporating all these variables.

Despite the fact that previous studies have not focused on test items in regards to typographic cues, achievement tests are still part of the educational materials that students are exposed to. Both weak and strong students may experience these signals differently as mentioned earlier (Narveson, 2001), even in tests.

The Influence of Demographics on Academic Achievement

An overarching goal of the NCLB act pushed to close the achievement gap between student groups (Ikegulu, 2004). Thus, a closer look at the different factors that influence student achievement may help educators close this gap. There are varying factors that may play a role in students' performance in standardized tests. A number of studies examined the influence of gender, ethnicity, SES or other demographic variables on student performance (Aguilar, 2010; Altshuler & Schmautz, 2006; Bedford, 2007; Chargois, 2008; Das, 2007; Heier, 2011; Ikegulu, 2004; Jaska, Hogan, & Wen, 2009; van Langen, Bosker, & Dekkers, 2006; Nichols, 2003; Quezada, 2008; Roane, 2008; Silva, 2010; Sirin, 2005; Slaughter, 2007; White, 1982). Each factor may play a different role, but a combination of all may be significant. In Quezada's (2008) study, for example, ethnicity (being the most influential), SES, and gender were all predictors of achievement in social studies tests. The sample of the study was eighth graders who took all the sections of the TAKS tests in an urban school district in South Texas in the year 2008.

For ESL students, the interaction between the language ability barrier and other factors such as ethnicity may result in lower performance rates. Roane (2008) found that this interaction is a significant obstacle to academic achievement in the content area of mathematics. Furthermore, Silva's (2010) study of 922 seniors who didn't graduate with their class due to TAKS failure in the years 2008, 2009, and 2010 at ten campuses in Texas showed that the interaction between gender, SES, and limited English proficiency was an indicator of student performance in all tests in general, and in the social studies test in specific. The majority of students who didn't pass the social studies TAKS test in the study and ended up not graduating were female (68% failed), LEP students (75% didn't pass) of lower SES (86% didn't pass). In Altshuler and Schmautz's (2006) study, the language ability barrier also interacted with ethnicity (in this case, Hispanic students) in student performance and test results.

Table 3 explains the percentage of students from different ethnic groups who passed the TAKS test in the years 2003, 2006, and 2009.

Table 3

TAKS Passing Rates (Percentages) For All TAKS Tests in 2003, 2006, and 2009 by Ethnicity.

	2003		2006		2009	
	<i>8th Grade</i>	<i>11th Grade</i>	<i>8th Grade</i>	<i>11th Grade</i>	<i>8th Grade</i>	<i>11th Grade</i>
State	69.9	49.8	58	66	67	76
African American	54.7	33.7	40	48	52	63
Hispanic	59.3	38.9	46	53	58	68
White	82.5	59.7	75	80	82	87
Native American	73.7	50.6	63	73	72	79
Asian/Pacific Islander	88.1	71.7	84	84	88	90

(TEA, 2003a; TEA, 2006a; TEA, 2009a)

Gender and Achievement

As mentioned earlier, one of the goals of the NCLB act is closing the achievement gap between student groups such as that between male and female students. Information on gender and gender gaps could be used by educators to develop instruction that is suitable for both genders (Quezada, 2008). Whether females or males perform better, and in what tests, is debatable depending on the type of study performed and the population. On mathematics tests, gender shows statistical significance especially when combined with other variables such as SES or ethnicity (Chudowsky & Chudowsky, 2010; Silva, 2010). Other studies such as those of Bedford (2007), Chargois (2008), Else-Quest, Hyde, and Linn (2010), and Slaughter (2007) (although with a limited sample of select 3rd grade students), indicated gender similarities in performance; meaning that achievement differences between genders were not significant.

In reading, females outperformed males across grade levels in Ikegulu's (2004) study with a sample drawn from 29 elementary and secondary schools in a school district in Texas. Other studies also found that females outperformed males in language arts and mathematics (Reiss, 2005; Quezada, 2008). On the contrary, in standardized science tests, female students are reported to have lower scores (e.g., Silva, 2010; Quezada, 2008). As for social studies, Silva's (2010) study of over 900 students who failed the TAKS tests shows that gender gaps were most evident in terms of test performance; 68% of females failed the social studies TAKS compared to 32% of the males. Earlier, Quezada (2008) found that males outperformed females in the social studies TAKS test. However, Bedford's (2007) study of 1,319 eighth grade students who took the 2005 TAKS tests in two school districts in South Texas showed no gender-based difference in civic knowledge.

Ethnicity and Achievement

The accountability system created as a result of the NCLB act of 2001 refers to student groups which particularly have been disadvantaged, especially in regards to their low income and their minority status (Ikegulu, 2004). As is the case with gender, there are achievement gaps based on ethnicity. The five ethnicities that are included in the TAKS Data File Format (TEA, 2006b) are: American Indian or Alaskan Native, Asian or Pacific Islander, African American, Hispanic, and White (not of Hispanic origin).

Generally, students from an ethnically minority background are reported to have higher failure rates in high-stakes tests (e.g., Silva, 2010; Quezada, 2008). Ethnicity may even have a stronger correlation with high school graduation than the ELL status (Flores

et al., 2012). In mathematics, the National Council of Teachers of Mathematics (2000) reports that there are achievement gaps based on ethnicity. Quezada (2008) echoes that by saying that mathematics, reading, science, and social studies' TAKS scores had ethnicity main effects in his study of eighth graders in schools led by either a principal with an elementary education teacher certification or a principal with a secondary education school certification. In social studies specifically, Quezada (2008) concluded that the strongest predictor of achievement was ethnicity. The study showed that non-Hispanic groups' scores were higher than those of the Hispanic group.

Students from minority ethnic groups (especially those in the ESL programs) are at a disadvantage for several reasons including lack of English language proficiency (Altshuler & Schmautz, 2006; Fry & Gonzales, 2008), having low education parents, and being raised in lower socio-economic environments, among others (Fry & Gonzales, 2008).

Social Economic Status (SES) and Achievement

The socio-economic status of students is one of the variables that have been examined extensively. Previous research (e.g., Das, 2007; Heier, 2011; Ikegulu, 2004; Jaska et al., 2009; Nichols, 2003; Quezada, 2008; Silva, 2010; Sirin, 2005; Slaughter, 2007; Stewart, 2009; White, 1982; Zamarripa, 2009) found evidence that the students' socio-economic status influences their achievement. In fact, SES, low SES more specifically, may have an even stronger effect on student success than other demographic variables (Ikegulu, 2004), although some may disagree in considering it a primary reason for the low performance of ELLs for example (Das, 2007). Economically

disadvantaged students are students who live in relatively low socio-economic environments, generally identified as poverty; as a result, based on their parents' income level, they are qualified to receive free or reduced fee lunch at school (TEA, 2009b).

The coding found in the TAKS Data File Format (TEA, 2006b) includes four categories for the SES variable: eligible for free meals under the National School Lunch and Child Nutrition Program, eligible for reduced-price meals under the same program, having other economic disadvantages, and not being identified as economically disadvantaged.

SES plays a role in student achievement. In schools that have a large percentage of low SES students, fewer students are found to have academic success (Stewart, 2009; Zamarripa, 2009). In a sample of student test scores from thirty seven Texas public high schools taken between the years 2003 and 2007, Jaska et al. (2009) reported that the TAKS scores were negatively affected by the percentage of economically disadvantaged students in the school district; as the number of disadvantaged students increased, the TAKS scores decreased. Further, in Heier's (2011) study of student performance in two schools that were labeled according to the funding they receive (Title I and non-Title I campuses), the difference in means in reading and mathematics scores were statistically significant. This was evident despite the finding of no significant difference when comparing the performance of economically disadvantaged students only between the schools.

Students who come from low SES backgrounds have a higher failure rate on high-stakes tests than others (Ikegulu, 2004; Silva, 2010; Nichols, 2003; Zamarripa,

2009). In relation to content area, in Quezada's (2008) study, the effect of SES favoring students who weren't classified as economically disadvantaged was evident in reading, mathematics, social studies, and science. Achievement on mathematics and social studies TAKS tests were found to be associated with the SES of Hispanic students, as SES was a statistically significant predictor of achievement in Aguilar's (2010) study. In regards to mathematics and science, achievement was positively related to students' SES; students with low SES scored lower than those of higher SES (van Langen et al., 2006). Moreover, Nichols (2003) found that, in each graduating class of the 6 different high schools in Indiana he examined between the years 2000 and 2002, lower SES students had greater failure rates on all the tests.

The socio-economic status of students may incorporate different aspects of their lives. Their parents' and families' needs and desires, the pressure they face in wanting to succeed, the stress of beginning working, not having peer or adult support, the parents' income, and others, are factors that affect ESL students' performance (Pyron, 2007). Nevertheless, understanding the interrelationship between SES and other demographic characteristics such as gender and ethnicity, and achievement, can help educators create better educational environments for our students in school (Chadwick & Harris, 2009).

Summary

As students receiving ESL services increase in number every year, their needs and challenges need to be addressed. The accountability system, in compliance with the NCLB act, puts ESL students under fire as they struggle to pass standardized state tests. The performance of ESL students on the TAKS test in general, and on the social studies

portion of it was shown to be influenced by the ESL students' limited English language abilities and their inadequate content knowledge, along with their demographic characteristics. The nature of the social studies TAKS test was explored extensively, with special focus on the five objectives of the test.

While educators are not supposed to teach to the test, question items on the test were explored in terms of their nature based on the objectives they fall in, and in terms of incorporating typographical signals. The appearance of the questions on the page as well as their content plays a role in student performance, as shown in previous research.

CHAPTER III

DESIGN OF THE STUDY

Theoretical Framework

This study used secondary data analysis on a large database provided by the State of Texas. Secondary data analysis uses data that has been collected by another party, other than the researcher him/herself. Thus, because the data for this study was collected by the Texas Education Agency (TEA) and then provided to the researcher, secondary data analysis, which is an analysis of “someone else’s data” (Smith, 2008, p.324), was utilized. Although there is no consensus on the exact definition of secondary data analysis, there are general definitions such as those described by Hakim (1983) and Jary and Jary (2000) that include the idea that it is the analysis of existing datasets which creates new knowledge. According to Strayhorn (2009), secondary data analysis is widely accepted in the realm of education. Secondary data can be either numeric or non-numeric such as interviews and ethnographic accounts (Smith, 2008). Numeric secondary data may include administrative data, data produced by national Census, and data gathered from tests of student performance. Of interest within secondary data analysis to this study is numeric secondary data since the dataset of study is numeric and character based. Although numeric secondary data analysis is not widely used in social sciences in general, and in education in specific, it does have its perks (Smith, 2008).

Secondary data analysis helps the researcher by saving time, effort, and money. Nothing much is required from the researcher, even in terms of personal and

professional traits (Glaser, 1963). It also provides a way in which social scientists can describe and test the social world around them, as well as contributing to issues such as accountability (Smith, 2008). Secondary data analysis can be combined with other approaches to maximize its usefulness especially in terms of generalizability (Smith, 2008). Specifically, large-scale secondary data analyses produce results that are robust with considerable generalization power; that is because the large-scale data consists of whole populations (Strayhorn, 2009). However, this generalizability power may lead to sacrificing a “degree of specificity” (Strayhorn, 2009, p.111).

Other drawbacks of secondary data analysis include missing data, complex sampling designs of how the data was originally collected (Strayhorn, 2009), challenges of preserving the original data (Glass, 1976), being erroneous, and being blamed for “reducing the complexity of social experiences to mere quantities” (Smith, 2008, p.336). However, for the purposes of this study, secondary data analysis provides the most benefit as the dataset included all the population and because there was no other way of obtaining such large-scale information other than the *secondary* way. As a side note, it is worthy to mention that the data used for this study could not be changed or influenced by the researcher in any possible way; thus, secondary data analysis was used.

The data provided was basically numeric and character based. Consequently, the quantitative method of analysis was chosen over the qualitative method. The quantitative method in the field of education has its benefits through generalizability for example, yet, like secondary data analysis, fails to recognize individual differences that cannot be quantified and overlooks specificity.

Finally, this study used non-experimental descriptive, correlational, and causal-comparative research designs to answer its research questions. As the research questions indicated, non-experimental design was required. In addition, a correlational research design was used to determine the relationships between variables especially in relation to the students' performance on the social studies portion of the TAKS. Finally, in order to investigate possible differences between groups, a causal-comparative research design was used.

Subjects of Study

The subjects of the current study were ESL students who were enrolled in grade 8 (April 2003 and April 2006) and Grade 11 (April 2006 and April 2009) and took the 8th and 11th grade social studies TAKS tests.

English as a Second Language (ESL) programs, provide some kind of support to ELLs in their native language while being instructed in English. The two forms of ESL programs include Content-Based Programs, and Pull-Out Programs. In the Content-Based Programs, a full-time certified teacher serves the students in providing supplementary instruction for all content areas. However, in the Pull-Out Programs, these services are provided by a part-time teacher who teaches only English language arts, while the students remain in the mainstream classroom for the other content areas (TEA, 2009e; TEA, 2012c). The number of ELL students who were enrolled in Bilingual Education programs in 2011 was 477,297, and those who were in ESL programs in the same year were 313,691 students (TEA, 2012c). The diversity in the State of Texas is evident in its student population with over 120 languages being

represented in Texas schools. Although 91% (744,949 in the year 2010) were Spanish speakers, other languages include Vietnamese (1.90%), Arabic (0.59%), Urdu (0.49%), and Korean (0.36%) (TEA, 2012c).

In the 2002-2003 school year, Texas public schools enrolled 572,019 students (13.5%) in Bilingual and ESL Education programs. The ethnic composition was 14.3% African Americans, 42.7% Hispanics, 39.8% Whites, 2.9% Asian/Pacific Islanders, and 0.3% Native Americans; economically disadvantaged students made up 51.9% (2,201,534 students) of the total student population (TEA, 2003a). However, the number of students in special programs for ELLs increased to 14.6% in 2006 (657,716 students) with the Hispanic population and the economically disadvantaged populations also increasing (45.3% and 55.6% respectively) (TEA, 2006a). Yet again, in 2009 enrollment in Bilingual and ESL Education programs showed an increase of about 2% from 2006 (757,146 students in total). The Hispanic, Native American, and Asian/Pacific Islanders populations also increased. Economically disadvantaged students made up 56.7% (2,681,474 students) of the student population for that year (TEA, 2009a).

The Academic Excellence Indicator System (AEIS) report produced by the TEA indicates that in the subgroup of economically disadvantaged students, passing rates in 2003 were 56.9% for 8th graders and 35.9% for 11th graders for all TAKS tests (TEA, 2003a). In 2006, the passing rates of economically disadvantaged students ranged from 44% in eighth grade, to 50% in eleventh grade (TEA, 2006a). Those passing rates increased in 2009 to 55% for eighth graders and 65% for eleventh graders (TEA, 2009a).

Data

The file released to the researcher by the TEA which contained student responses on the social studies TAKS test scores was the major data source for this study. The TAKS social studies passing percentages and item response analysis for select students were collected from the TEA as it is information that is available upon request.

An additional source of data was the released tests of the social studies TAKS. These released tests are available online on the TEA website. The question items were analyzed to determine which fell under the excerpt-based question items and which were bullet-point question items, in order to address the third research question. Finally, Data File Format with Student Item Analysis documents were used as part of understanding the data and the coding of the TEA. These data files are also provided via the TEA website.

The data, which included the test scores of students in Grade 8 who took the social studies TAKS test in April 2003 and April 2006, and students in Grade 11 who took the test in April 2006 and April 2009, was obtained from the Texas Education Agency (TEA) as a response to an official letter of request sent by the researcher. The data provided by the TEA included a wide range of information on students in both 8th and 11th grade who were given the TAKS assessment in the years 2003, 2006, and 2009. The data was coded; it included the administration years, grade designation, student identification number, as well as demographic information such gender, ethnic classification, economically disadvantaged classification, participation in the Title I Program, and migrant status. The researcher was interested in the SES of the students,

and chose to use only the economically disadvantaged classification as a measure of SES.

In addition, information on the program students were participating in was included such as participation in a Bilingual program, ESL program, gifted/talented program, and a special education program. Only the scores of students who were enrolled in the ESL program were selected for inclusion in the study.

In relation to social studies, the test language information, oral administration, waivers (for example, foreign exchange), and LEP postponement were provided in order to refine the data further as the researcher was interested in a homogenous group which was a representative of the ESL student who was given the test in its written form in the years mentioned above. In addition, social studies scores given by objective (each of the five objectives reported separately), and total social studies raw scores, along with social studies scale scores, 'met standard' classification, and commended performance in the area of social studies were provided. Finally, social studies objective numbers, item correct responses, and item student responses were given.

The information provided to the researcher was in the form of comma-delimited text data files, which were imported into Microsoft Excel and reorganized (see Figure 1 below).

```

STUDENTI, YEAR, GRADE, SEX, ETHNIC, DISADV, TITLEI, MIGSTA, LEPROF, BILING, ESLANG, SPECED, GIFTAL, ATRISK, VOCE
*5F63LB26,06,08,F,4,0,6,0,C,0,1,0,0,1,0,E,,0,0,0,,,,,S,,4,4,6,4,4,22,2061,0,0,0,0,030403050202020
K**61LL42,06,08,M,4,1,6,0,C,0,1,0,0,1,0,E,,0,0,0,,,,,S,,3,2,5,5,2,17,1994,0,0,0,0,030403050202020
1F173LA22,06,08,F,4,1,6,0,C,0,1,0,0,1,0,E,,0,0,0,,,,,S,,6,3,3,5,3,20,2035,0,0,0,0,030403050202020
A4*08*KL2,06,08,M,4,0,6,0,0,0,0,0,0,0,1,0,E,,0,0,0,,,,,S,,4,3,5,4,4,20,2035,0,0,0,0,030403050202020
A0362LF24,06,08,F,4,2,6,0,0,0,0,0,0,0,1,0,E,,0,0,0,,,,,S,,7,3,8,7,5,30,2164,1,0,0,0,030403050202020
A5L5FLE21,06,08,M,4,9,6,0,C,0,1,0,0,1,1,E,,0,0,0,,,,,S,,3,4,2,3,0,12,1919,0,0,0,0,030403050202020
*2150LB25,06,08,F,4,0,6,0,0,0,0,0,0,0,1,0,E,,0,0,0,,,,,S,,5,4,6,6,8,29,2151,1,0,0,0,030403050202020
BADSTUNUM,06,08,M,5,1,6,0,0,0,0,0,0,0,1,0,E,,0,0,0,,,,,S,,10,5,5,6,6,32,2192,1,0,0,0,030403050202020
30E63LB1F,06,08,F,4,0,6,0,0,0,0,0,0,0,1,0,E,,0,0,0,,,,,S,,10,4,4,6,7,31,2178,1,0,0,0,030403050202020
A1*54L120,06,08,F,4,9,6,0,0,0,0,0,0,0,1,1,E,,0,0,0,,,,,S,,9,6,8,9,7,39,2309,1,0,0,0,030403050202020
SA25F80KE,06,08,F,4,9,6,0,0,0,0,0,0,0,1,1,E,,0,0,0,,,,,S,,9,5,7,6,5,32,2192,1,0,0,0,030403050202020
*3354L*22,06,08,M,4,9,6,0,C,0,1,1,0,1,1,E,,0,0,0,,,,,S,,11,6,5,8,6,36,2253,1,0,0,0,030403050202020
S4L478130,06,08,F,4,1,6,0,C,0,1,0,0,1,0,E,,0,0,0,,,,,S,,2,3,7,6,2,20,2035,0,0,0,0,030403050202020
SF**78024,06,08,M,4,9,6,0,0,0,0,0,0,0,1,0,E,,0,0,0,,,,,S,,7,5,8,7,6,33,2206,1,0,0,0,030403050202020
*1257L12E,06,08,M,4,9,6,0,0,0,0,0,0,1,0,1,0,E,,0,0,0,,,,,T,X,,0,0,0,0,0,0,1361,0,0,0,1,030403050202020
*E364LK24,06,08,M,4,9,6,0,0,0,0,0,0,0,0,0,E,,0,0,0,,,,,S,,12,5,9,12,7,45,2493,1,1,0,0,030403050202020
A**E68LK25,06,08,M,4,9,6,0,C,0,1,1,0,1,0,E,,0,0,0,,,,,T,X,,0,0,0,0,0,0,1361,0,0,0,1,030403050202020
SAL558EK3,06,08,F,4,9,6,0,C,0,1,0,0,1,1,E,,0,0,0,,,,,S,,2,2,2,4,1,11,1901,0,0,0,0,030403050202020
A3E52LV23,06,08,F,4,9,6,0,0,0,0,0,0,0,1,0,E,,0,0,0,,,,,S,,5,4,6,4,7,26,2112,1,0,0,0,030403050202020
A4460L*26,06,08,F,4,9,6,0,C,0,1,1,0,1,0,E,,0,0,0,,,,,T,X,,0,0,0,0,0,0,1361,0,0,0,1,030403050202020
SAE5F8EK5,06,08,M,4,0,6,0,0,0,0,0,0,0,1,0,E,,0,0,0,,,,,S,,11,6,8,6,7,38,2289,1,0,0,0,030403050202020
A4F60L*24,06,08,F,4,9,6,0,0,0,0,0,0,0,1,0,E,,0,0,0,,,,,S,,11,3,8,9,7,38,2289,1,0,0,0,030403050202020
*5567LL16,06,08,M,4,9,6,0,0,0,0,0,0,0,1,0,E,,0,0,0,,,,,S,,12,3,9,10,8,42,2400,1,1,0,0,030403050202020
S41478132,06,08,F,4,1,6,0,C,0,1,0,0,1,0,E,,0,0,0,,,,,S,,2,3,6,6,3,20,2035,0,0,0,0,030403050202020
A4464LK33,06,08,F,4,9,6,0,C,0,1,1,0,1,0,E,,0,0,0,,,,,T,X,,0,0,0,0,0,0,1361,0,0,0,1,030403050202020
*0E6FLO27,06,08,F,4,9,6,0,C,0,1,1,0,1,0,E,,0,0,0,,,,,T,X,,0,0,0,0,0,0,1361,0,0,0,1,030403050202020
1FE72LB1F,06,08,M,4,9,6,0,C,0,1,1,0,1,0,E,,0,0,0,,,,,T,X,,0,0,0,0,0,0,1361,0,0,0,1,030403050202020
S4L468133,06,08,M,4,9,6,0,C,0,1,0,0,1,0,E,,0,0,0,,,,,S,,8,4,8,6,8,34,2221,1,0,0,0,030403050202020

```

Figure 1. An Image of the Spring 2006 Grade 8 Social Studies TAKS Test Data Received from the Texas Education Agency.

After that, data were further disaggregated and organized to adhere to the purposes of this study. Only data of students participating in the ESL program (in both its formats: Pull-Out, and Content-Based) with a score in the social studies regular TAKS test were chosen for this study. The data of this group were refined further to exclude information about students who took the oral version of the test, students who provided no responses to the questions on the test, students whose IDs were flagged and not provided, foreign exchange students, students who took any other version of the TAKS test (such as the TAKS versions designed to accommodate special education students), students who needed any kind of test accommodations, and students who were eligible for LEP postponement. Information of students who were masked (as per the TEA requirement of masking certain student information as part of complying with the

Family Educational Rights and Privacy Act (FERPA)) was also excluded. The following listed items represent the columns of data included in this study.

- Alternate Student Identification Number
- Administration Year
- Grade
- Sex
- Ethnic Classification
- Economically Disadvantaged Status
- Participation in an ESL Program
- Social Studies Objective 1 (U.S. History) Scores
- Social Studies Objective 2 (Geography) Scores
- Social Studies Objective 3 (Economics) Scores
- Social Studies Objective 4 (Civic Knowledge) Scores
- Social Studies Objective 5 (Social Studies Skills) Scores
- Social Studies Raw Score
- Social Studies Met Standard
- Social Studies Item Objective Numbers
- Social Studies Item Correct Responses
- Social Studies Item Student Responses

The number of questions on the social studies TAKS for both the 8th and 11th grades was consistent in the years of interest (2003, 2006, 2009). The Grade 8 social studies TAKS test consisted of 48 questions, and the Grade 11 social studies TAKS test consisted of 55 questions. The social studies item responses of the students were dissimilated into columns and matched with an objective number according to the information provided in the social studies item objective numbers that correspond to each year's test items (see Figure 2 below).

6 obj-2	7 obj-2	8 obj-5	9 obj-4	10 obj-1	11 obj-3	12 obj-4	13 obj-3	14 obj-2	15 obj-1
+	+	+	+	+	+	H	+	+	+
H	+	+	C	J	C	+	+	+	+
+	+	+	A	H	+	+	+	F	+
+	+	F	A	+	+	J	B	+	A
+	+	G	+	H	B	+	B	+	C
+	C	+	B	H	+	+	+	+	+
F	B	+	C	+	+	+	+	+	C
H	+	+	+	J	+	+	+	F	D
J	+	+	A	+	D	+	+	+	A
+	+	+	+	J	+	F	C	+	D
J	+	+	+	+	+	+	+	+	A
+	+	+	C	H	+	J	B	F	D
H	+	+	+	+	+	H	C	+	C
H	+	J	B	J	B	+	+	+	C
H	D	G	+	+	+	+	+	+	C
F	D	+	A	+	C	J	+	F	C
H	+	J	C	+	D	F	B	H	A

Figure 2. Sample of Separating Student Responses and Matching Objectives to Question Items (Spring 2006, 8th grade questions).

For the purpose of the different research questions in this study, separate spread sheets were created in Excel to include only the columns that provided data needed to address each question. Grade levels were combined in some, and others included the whole student population that fit the inclusion criteria.

Instrumentation

The Texas Assessment of Knowledge and Skills (TAKS) assessment was used in this study to measure the academic performance of students in the content area of social studies. The test consists of multiple choice question items. Each multiple-choice question has four possible responses. In order to pass the social studies part of the TAKS assessment, a student should answer a number of questions correctly to have met the standard level. This number is different depending on the year the test was administered

and the grade level. For the 8th grade, out of 48 questions, in Spring 2003, 19 correct answers were needed; however, in 2006 and 2009, 25 correct answers were required to have met the standard level. For 11th grade, in April 2003, 22 correct answers out of 55 were required to pass, and in 2006 and 2009, 21 correct answers out of 55 were needed to successfully pass this section of the TAKS (TEA, 2011c). The reason for that is that the State Board of Education (SBOE) adopted the TAKS scale scores instead of the raw scores in order for tests to be equally difficult across administrations starting September 2003 (TEA, 2011c). Since the TAKS tests may differ in difficulty across administrations, the raw score cuts may be different for each test administration, whereas the scale score standards remain constant (TEA, 2011c).

Consistency, stability, and precision of test scores all represent reliability (Gall, Gall, & Borg, 2003). Reliability pertains that the same test generates consistent results in regards to student performance. The TAKS assessment measures reliability based on one test administration, which is known as an internal consistency measure. The Kuder-Richardson 20 (KR20) is a type of internal consistency measure that is used to estimate the reliability of Texas assessments with only multiple choice items (TEA, 2010). Another type of measurement is used for mixed-type tests with both multiple choice and constructed response items; it is called the Stratified Coefficient Alpha, as well as the inter-reliability correlation measurement in other versions of the test such as the Alternative TAKS. For the purposes of this study, the researcher was interested in the KR20, which estimates the reliability of the test as the ratio of true score (score with no measurement error) variance to observed score variance.

Tables 4 and 5 report the internal consistency values for the 8th and 11th grade TAKS social studies subject area test by objective for the years 2003, 2006, and 2009. In addition, the KR20 reliability estimates for the social studies TAKS test by gender and ethnicity for the years 2003, 2006, and 2009 are presented in Table 6. All three tables show relatively high scores of KR20, indicating that they are good, and in some cases excellent.

Table 4

Internal Consistency (KR20 Reliability) for Grade 8 Social Studies Exam in 2003, 2006, and 2009.

	2003	2006	2009
Complete Subset	0.884	0.913	0.904
Issues and Events in U.S. History	0.710	0.773	0.745
Geographic Influences on History	0.536	0.606	0.538
Economic and Social Influences on History	0.564	0.653	0.628
Political Influences on History	0.654	0.688	0.706
Critical Thinking Skills	0.491	0.668	0.560

(TEA, 2003b; TEA, 2006c; TEA, 2009d)

Table 5

Internal Consistency (KR20 Reliability) for Grades 11 Social Studies Exam in 2003, 2006, and 2009.

	2003	2006	2009
Complete Subset	0.906	0.903	0.900
Issues and Events in U.S. History	0.697	0.689	0.722
Geographic Influences on History	0.535	0.558	0.587
Economic and Social Influences on History	0.674	0.641	0.663
Political Influences on History	0.669	0.673	0.608
Critical Thinking Skills	0.706	0.720	0.618

(TEA, 2003b; TEA, 2006c; TEA, 2009d)

Table 6*KR20 Reliability Estimates for the Social Studies TAKS Test in 2003, 2006, and 2009.*

	2003		2006		2009	
	Grade 8	Grade 11	Grade 8	Grade 11	Grade 8	Grade 11
Total Group	0.884	0.906	0.913	0.903	0.904	0.900
Female	0.870	0.888	0.904	0.897	0.896	0.894
Male	0.895	0.920	0.921	0.906	0.911	0.903
African American	0.852	0.879	0.899	0.886	0.894	0.889
Asian	0.892	0.907	0.903	0.900	0.884	0.901
Hispanic	0.862	0.884	0.904	0.896	0.899	0.897
Native American	0.863	0.900	0.906	0.876	0.895	0.886
White	0.881	0.906	0.904	0.885	0.892	0.879

Retrieved from the following appendices in the 2002-2003, 2005-2006, and 2008-2009 TEA Technical Digests: Appendix 7, 2003; Appendix 26, 2006; Appendix B, 2009.

Related to reliability, validity is important in order to determine if the test measures what was intended to be measured. The TAKS test is intended to measure students' knowledge and understanding of the TEKS; thus, the measure needs to test those to ensure correct inferences about student achievement. The inferences made from the test scores are to be appropriate, meaningful, and useful in order to ensure validity (Gall et al., 2003). Different types of validity evidence are used with the TAKS test: evidence based on test content, evidence based on response processes, evidence based on internal structure, evidence based on relations to other variables, and evidence based on consequences of testing (TEA, 2010). However, that support of validity is an ongoing process.

Because the TAKS test is standards-referenced, it is based on an extensive explanation of the content in which it measures (TEA, 2006c). Evidence based on test

content is basically relating the test content to the TEKS, which is accomplished during the test development process. This process includes a review of the TEKS, a preparation of test objectives and how to assess them, a process of item development based on a test blueprint, a review of the items, field-testing of the items, among many others, which are conducted by Texas educators, the TEA staff members, and university-level experts (TEA, 2010). Additionally, the content is reviewed for any bias that may exist. Test items are examined for potential ethnic bias, and for possible impact on minority groups (TEA, 2009d). In the development of the TAKS tests, educator input, which include annual revisions of proposed test items before and after field-testing, was used (TEA, 2004a). In addition, experts in different content areas met yearly to review the content of each of the high school level TAKS tests; this is referred to as a content validation review to ensure that these tests are of the highest quality possible (TEA, 2004a).

For the evidence based on response processes, both theoretical and empirical evidence is gathered to show that response processes support the accurate measurement of the construct. Evidence based on internal structures is evaluated for all the students and estimated for the full test and for each objective within a content area, using the KR20 or the Stratified Coefficient Alpha depending on the type of question items, to ensure the conformity of the test parts to the construct of the test (TEA, 2010). Another type of validity evidence is based on relationships between test performance and performance on other measures, also called criterion-related validity. Annual grade correlation studies were conducted to provide evidence that support the validity of the test (TEA, 2010). Finally, to collect evidence based on consequences of testing (also

known as consequential validity), surveys were given to educators and other assessment program stakeholders.

Analytic Framework

This study utilized a quantitative approach to analysis since the information provided by TEA was in the form of interval data. The data was refined and coded by the researcher, then entered into EXCEL sheets for use with the computer program of statistical analysis JMP. Initially, descriptive statistics were conducted on the demographic data. To examine the first question, an ANOVA test was conducted with a null hypothesis of all means of the objectives being equal, followed by Tukey's method of pairwise comparison. After identifying the students who remained in the ESL program, a multivariate paired comparison (MANOVA) was conducted to address the second question. The third question, which involved the two types of test questions of interest, was analyzed using a directional independent t-test. A paired t-test was used to address the fourth question to identify a systematic pattern in relation to the excerpt-based question items examined in the previous research question. Finally, the last question was addressed using a multiple regression analysis; the independent variables being gender, ethnic classification, and SES, and the dependent variable being the academic performance measured by the social studies section of the TAKS assessment (the raw score).

Participants

The entire population of ESL students whose data met the criteria set by the researcher, rather than sampling from each grade or each year, was chosen for study.

Students who participated in an ESL program in 8th grade in the years 2003 and 2006, and students who were part of an ESL program in 11th grade in the years 2006 and 2009 in the State of Texas, and those who received a score in the social studies TAKS exam were included in this study. Those who took alternative forms of the test were excluded, and those who needed accommodations or were exempt for whatever reason were eliminated as well.

A total of 40,568 students' data were used for this study. The total is a combination of data of students who took the test in 2003 in 8th grade (9751 students), and those who took the test in 2009 in 11th grade (9047 students). The researcher also used the data of students in both 8th and 11th grades for the year 2006 (13813 students and 7957 students, respectively) for a total of 21770 student data in 2006. As for matching the data of students who took the social studies TAKS and were promoted from 8th grade in 2003 to 11th grade in 2006, 1,706 ESL students were found. However, these numbers increased for the data set matching students who took the test in 8th grade of 2006 then promoted to 11th grade in 2009 to 3167 students. The total number of students' data that was matched between 8th and 11th grades was 4,873 students. Table 7, Table 8, and Table 9 display the breakdown of the participants' data in 2003, 2006, and 2009 by various categories.

Table 7

Gender, Ethnicity, and SES of Participants in the ESL Program Completing the Social Studies TAKS Test in Total.

	Grade 8	Grade 11
Total	23,564	17,004
Gender:		
Male	12,661 (53.7%)	8,625 (50.7%)
Female	10,903 (46.3%)	8,379 (49.3%)
Ethnicity:		
American Indian or Alaskan Native	0 (0.0%)	0 (0.0%)
Asian or Pacific Islander	762 (3.2%)	944 (5.6%)
African American	139 (0.6%)	142 (0.8%)
Hispanic	22,470 (95.4%)	15,813 (93.0%)
White	193 (0.8%)	105 (0.6%)
Social Economic Status:		
Eligible for free meals	16,486	10,638
Eligible for reduced-price meals	1,434	1,060
Other economic disadvantages	3,082	2,311
Total economically disadvantaged	21,002 (89.1%)	14,009 (82.4%)
Not identified as economically disadvantaged	2,562 (10.9%)	2,995 (17.6%)

Table 8

Gender, Ethnicity, and ESL Status of 8th Grade Participants in the ESL Program Completing the Social Studies TAKS Test in Spring 2003 and Spring 2006.

	Spring 2003	Spring 2006
Total	9,751	13,813
Gender:		
Male	5,325 (54.6%)	7,336 (53.1%)
Female	4,426 (45.4%)	6,477 (46.9%)
Ethnicity:		
American Indian or Alaskan Native	0 (0.0%)	0 (0.0%)
Asian or Pacific Islander	399 (4.1%)	363 (2.6%)
African American	50 (0.5%)	89 (0.7%)
Hispanic	9,204 (94.4%)	13,266 (96.0%)
White	98 (1.0%)	95 (0.7%)
Social Economic Status:		
Eligible for free meals	7,117	9,369
Eligible for reduced-price meals	530	904
Other economic disadvantages	906	2,176
Total economically disadvantaged	8,553 (87.7)	12,449 (90.1%)
Not identified as economically disadvantaged	1,198 (12.3%)	1,364 (9.9%)

Table 9

Gender, Ethnicity, and ESL Status of 11th Grade Participants in the ESL Program Completing the Social Studies TAKS Test in Spring 2006 and Spring 2009.

	Spring 2006	Spring 2009
Total	7,957	9,047
Gender:		
Male	4,043 (50.8%)	4,582 (50.6%)
Female	3,914 (49.2%)	4,465 (49.4%)
Ethnicity:		
American Indian or Alaskan Native	0 (0.0%)	0 (0.0%)
Asian or Pacific Islander	434 (5.5%)	510 (5.6%)
African American	64 (0.8%)	78 (0.9%)
Hispanic	7,393 (92.9%)	8,420 (93.1%)
White	66 (0.8%)	39 (0.4%)
Social Economic Status:		
Eligible for free meals	5,075	5,563
Eligible for reduced-price meals	472	588
Other economic disadvantages	1,034	1,277
Total economically disadvantaged	6,581 (82.7%)	7,428 (82.1%)
Not identified as economically disadvantaged	1,376 (17.3%)	1,619 (17.9%)

Organization of TAKS Test Question Items

For each of the four social studies TAKS tests analyzed in this study (Grade 8 in the years 2003 and 2006; Grade 11 in the years 2006 and 2009), the researcher identified question items that fell within the categories of excerpt-based questions and bullet-point questions. An excerpt-based question item is an item which contains a box with written text that includes an indication to the source of information prior to or following the text. The item is introduced by the following instruction sentence provided in the test booklet, “Use the excerpt and your knowledge of social studies to answer the following question.” A question item is considered a bullet-point question if it contains enboxed written text organized in bullet points or using enumeration (including at least one piece of information); the text is organized in points. The item is introduced by the following instruction sentence provided in the test booklet, “Use the information and your knowledge of social studies to answer the following question.” Figures 3 and 4 show samples of each type of question identified for study in this research.

Use the excerpt and your knowledge of social studies to answer the following question.

Excerpt from Washington's Farewell Address

The great rule of conduct for us in regard to foreign nations is in extending our commercial relations, to have with them as little political connection as possible. . . . It is our true policy to steer clear of permanent alliances with any portion of the foreign world. . . .

- 34 Which of the following best reflects the point of view expressed by President George Washington in the excerpt above?
- F Foreign nations are a military threat.
 - G Permanent alliances are part of a nation's foreign policy.
 - H Foreign nations profit from importing goods.
 - J Permanent alliances with foreign nations should be avoided.

Figure 3. Sample Excerpt-Based Question item from the 8th grade Social studies Released TAKS Exam for the Year 2006.

Use the information in the box and your knowledge of social studies to answer the following question.

- Small tracts of land
- Family-based workforce
- Crops mainly grown for personal consumption

20 What type of economy is described by the information above?

- F Subsistence agriculture
- G Cottage industry
- H Commercial agriculture
- J Command economy

Use the information in the box and your knowledge of social studies to answer the following question.

- Outlawed discrimination in hotels, motels, restaurants, theaters, and all other public accommodations engaged in interstate commerce
- Prohibited discrimination by government agencies that receive federal funding
- Prohibited discrimination in hiring on the basis of race, color, religion, sex, or national origin
- Established the Equal Employment Opportunity Commission

21 Which of the following is described by the list above?

- A The Voting Rights Act of 1965
- B *Brown v. Board of Education*
- C The Civil Rights Act of 1964
- D *Regents of the University of California v. Bakke*

Figure 4. Sample Bullet-Point Question Items from the 11th Grade Social Studies Released TAKS Exam for the Year 2009.

To ensure inter-rater reliability, a common practice among researchers to evaluate the quality of data (Gwet, 2008), the researcher enlisted two individuals to identify the two types of question items of interest to this study in the released social studies TAKS tests for the years 2003, 2006, and 2009. Inter-rater reliability is concerned with the closeness of the scores that are assigned by a group of raters to the same study data; the closer the scores, the higher the reliability of the data (Gwet, 2008). A pilot was conducted to check for any discrepancies at an early stage of the analysis, in order for it to be addressed through alteration of definition or enhancement of training.

The two individuals, each with a Master's degree in their respective fields, were given a brief training by the researcher to help them identify the two types of questions of interest to the study. The definitions of both an excerpt-based question and a bullet-point question were explained and sample question items were presented. The 2004 social studies TAKS tests for 8th and 11th grades were given to the two raters as a pilot. The data used for this study used the social studies TAKS tests for 8th grade in the years 2003 and 2006, and for the 11th grade in the years 2006 and 2009, so those four tests were provided to the two raters in PDF format. They were then asked to identify excerpt-based and bullet-point question items in each.

The two raters and the researcher were in 100% agreement on the categorization of the questions. Twenty six question items were identified for each of the categories across the years and grades analyzed. On the grade 8 social studies TAKS test in 2003, 4 questions were categorized as excerpt-based, and one question was bullet-point. However, more question items were found for the same grade level in the year 2006, where 10 question items were under the excerpt-based category, and 7 were bullet-point questions. As for the 11th grade social studies TAKS tests, in the year 2006, seven questions were excerpt-based and 6 were bullet-point questions. Finally, in the year 2009, the 11th grade TAKS test yielded 5 excerpt-based questions, and 12 bullet-point questions.

Analytic Process

Because the data received from TEA was in the form of comma-delimited text files, as mentioned earlier, the researcher first manipulated it to be able to import the

data to Excel sheets for analysis. Therefore, the data was first imported to EXCEL into separate columns and rows, to make recognizing the categories easier. Then, all the mask-flagged data was eliminated; some students' results data were masked because fewer than five students were tested in a "Y" masking variable, in order to comply with the Family Educational Rights and Privacy Act (FERPA) (Jennifer J. Eaton, personal communication, June 30, 2011). Other columns were deleted from the data set in order to have the data only of the students who participated in an ESL program with a score in social studies TAKS tests for the years 2003, 2006, and 2009. Consequently, students with no score, students who took different versions of the test, students who needed test accommodations, and others who were eligible for test exclusion or postponement were eliminated from this study. The data of a homogenous group was the aim of the refinement and organization of the data provided by the TEA.

The researcher then separated the string of values within the Key column and the Student Response column, into separate cells in order to conduct the analysis by question. Each question was matched with an objective number (1 through 5) according to the information provided by the TEA in the text files they sent the researcher. Each question was given a question number and an objective number. The correct answer was indicated with a "+" sign and the wrong answer was displayed by a letter (A, B, C, D...etc.) that represented the wrong choice. The researcher converted the "+" signs to the number "1", and each incorrect answer (regardless of the letter chosen) to a "0". All question columns contained either a "1" or a "0" representing correct and incorrect answers respectively.

Descriptive statistics were conducted on demographic data as a general overlay of the analysis in order to describe the sample. To answer the first research question, an ANOVA was conducted with a null hypothesis of equal means of the five objectives. However, before doing that, the scores (number of correct answers) in each objective were converted to a percentage rather than a raw score; that is because the number of questions is unequal for each objective. To determine which objective is the most difficult and the rank of objectives in terms of their difficulty, a pairwise comparison using Tukey's Honestly Significant Difference (HSD) test was conducted in order to control Type I errors.

Before addressing the second and fourth research questions that were concerned with showing if a systematic difference existed throughout the years, the data needed to be reorganized. The data requested from the TEA included student identification numbers. The TEA provided made-up ID numbers to protect the identity of the students. Using these IDs, the researcher matched the raw results of ESL students in 8th grade, who took the test in the years 2003 and 2006, with the same IDs of students in 11th grade who remained in an ESL program and took the test in the years 2006 and 2009 using the VLOOKUP function in Excel. Thereafter, the sample was used to answer the second and fourth research questions. After obtaining the sample of matched IDs of the students who remained in the ESL program, a MANOVA was conducted to address the second research question with two dependent variables (scores from 8th grade, and scores from 11th grade).

The third and fourth research questions, which addressed the question item types, required several steps prior to the analysis of data. A definition of the types of question items of interest needed to be identified. Based on the definitions, classification of the question items into the two question item types (excerpt-based and bullet-point questions) was done by the researcher for each released test for the years 2003, 2006, and 2009 after a pilot using the 2004 exam. Two other volunteers also classified the questions in the released tests according to training provided by the researcher to ensure inter-rater reliability.

Addressing the third research question involved testing whether there was a difference in the correct response rate for each question item within the two question types ('excerpt' and bullet-point format). To do that, the percentage of correct answers for each question (identified as one of the two question types of interest to this research) was calculated. So, after identifying which questions were excerpt-based and which were bulleted, an indication of the individual answers of students (1 if answered correctly, and 0 if failed to answer the question correctly) was done. Then, a percentage of correct answers was calculated for each question. Since, after initial analysis, it was found that an equal number of excerpt-based questions and bullet-point questions existed in the years examined, a directional independent t-test of equal sample sizes of the correct percentages was conducted to answer the third research question. The null hypothesis of the test was that the two means of % correct in expert-based questions is equal to the % correct in bulleted questions.

The fourth research question was analyzed using a paired t-test. The procedures remained in the ESL program were identified (using the VLOOKUP function in EXCEL to match the student IDs). Then, as done in RQ3, the excerpt-based question items were identified in the released tests, but the percentages of correct answers were calculated for each student (unlike RQ3) only in the excerpt-based questions. So for each student, a percentage of correct answers on excerpt-based questions was calculated for 8th grade, then again for 11th grade. Then, a paired t-test was conducted to compare the students' performance throughout the years (in 8th grade and later in 11th grade).

Finally, to address the last research question, a multiple regression analysis was used using a numeric dependent variable (percentage of students' raw scores in social studies- meaning the number of correct answers). The independent variables were gender, ethnic classification, and SES. Each independent variable was addressed separately.

As a final note, all the assumptions related to each statistical test were verified to be met in order to ensure the effectiveness of the tests. The minimum level of significance that was accepted in the statistical tests was $P < 0.05$. When the assumptions were not met, other non-parametric methods were used to ensure the robustness of the tests and the significance of the interpretations they yielded.

CHAPTER IV

PRESENTATION AND ANALYSIS OF DATA

The presentation and analysis of data chapter is arranged into five sections that correspond with the five major research questions. The first section focuses on the first research question centered on comparing the performance of ESL students in the five objectives of the social studies TAKS. The next section focuses on the second research question which is similarly concerned with comparing that performance in a systematic way throughout the years. The third section, which represents the third research question, is centered on comparing the percentages of correct responses on item formats which include reading texts (excerpt-based and bullet-point questions). In relation to that, the fourth section, corresponding with the fourth research question, is concerned with the performance of students who remained in the ESL program on excerpt-based questions, throughout the years. Finally, the last section focuses on the fifth research question as it analyzes the effects of demographics (gender, ethnic classification, and SES) on ESL students' performance on the social studies TAKS exam.

Research Question 1: Analysis of Performance by Objective

To answer the first research question, “Is there a significant difference in the performance of ESL students across the five objectives of the social studies TAKS exam?” a one-way ANOVA was conducted. Although ANOVA is used for examining “the relative contribution of different sources of variation (factors or combination of factors, *i.e.* the predictor variables) to the total amount of variability in the response

variable” (Quinn & Keough, 2002, p. 195), ANOVA is utilized here to test the null hypothesis of population group means are equal. Before testing the hypothesis, descriptive statistics are presented to show the number, mean, standard deviation, and confidence intervals for the mean by objective for all five objectives in the social studies TAKS tests (see Table 10).

Table 10

Descriptive Statistics for the Performance of ESL students in the Social Studies TAKS in the Five Objectives.

	Mean*	Standard Deviation	95% Confidence Interval for Mean
Objective 1: History	0.48778	0.21170	0.48558-0.48998
Objective 2: Geography	0.55587	0.25260	0.55367-0.55807
Objective 3: Economics	0.56860	0.21549	0.56640-0.57081
Objective 4: Politics	0.51112	0.21760	0.50892-0.51332
Objective 5: Social Studies Skills	0.56758	0.23106	0.56538-0.56978

* Mean of % correct responses for each objective.
N for each objective= 40568.

The significance values for the one-way ANOVA used to analyze the percentages of correct answers in each of the five objectives of the social studies TAKS as performed by ESL students are reported in Table 11.

Table 11

Summary of One-Way ANOVA to Analyze the Performance of ESL Students in Social Studies TAKS Based on Test Objectives.

Source	<i>df</i>	<i>SS</i>	<i>MS</i>	F	p
Objectives	4	218.072	54.5181	1065.628	<.0001
	202835	10377.147	0.0512		
	202839	10595.220			

The assumptions of ANOVA were checked as follows. First, the measurement of the dependent variable, which is the percentages of the correct responses for each objective, is a ratio as it represents the number of correct answers divided by the total number of questions in each objective; this meets the assumption of the dependent variable being a ratio. Second, the independence of data observations of one another has been met per the manner in which the data was obtained; each student took the social studies TAKS exam one time in the spring for the specific years mentioned.

As for the normality assumption, the data was found not to satisfy the normality assumption using a Kolmogorov-Smirnov (KSL) Goodness-of-Fit Test of the residuals. The KSL Goodness-of-Fit Test resulted in a probability of 0.0100 which indicates that the null hypothesis of data being from a normal distribution is rejected at the $p=0.05$ level because $0.01 < 0.05$.

Given that the first two assumptions were met and that the sample size was large ($N= 40568$) that a violation of the normal distribution of data would not cause major problems according to Quinn and Keough (2002), the researcher used a Levene statistic to test for the homogeneity of variances. The Levene test resulted in a small p-value

(Prob>F = <.0001) rejecting the null hypothesis of equal variances. Consequently, there was a violation of the assumption of homogeneity of variances, although Quinn & Keough (2002) suggest that if the sample size is large, “these tests could reject the H_0 of equal variances when the ANOVA F test would still be reliable” (p. 205). According to Pallant (2007), the Welch robust test of equality of means should be consulted anytime a violation of homogeneity of variances occurs. Thus, as measures of precaution, Welch’s ANOVA Test, which allows for unequal standard deviations, is reported. An F Ratio of 1143.1195 with 4 degrees of freedom has a prob>F of <.0001 which is a small p-value that indicates rejecting the null of equal means. Therefore, both the F Ratio of ANOVA ($F(4, 202835) = 1065.628, p = <.0001$) as indicated in Table 11, and the F ratio and p-value of the Welch’s test ($F(4, 101324) = 1143.1195, p = <.0001$, refer to Table 12) show that the means of the five objectives are unequal. That is, given that there were no outliers found and that the independence assumption is satisfied as mentioned earlier.

Table 12

Summary of the Welch Test to Analyze the Performance of ESL Students in Social Studies TAKS Based on Test Objectives.

F Ratio	DFNum	DFDen	Prob>F
1143.1195	4	101324	<.0001

Furthermore, Tukey-Kramer HSD Test of pair-wise comparison shows that all the objective means are statistically significantly different from each other (with a p-value of <.0001), with the exception of objective 3 (economics) and objective 5 (social

studies skills). In the comparison of objective 3, economics (M=0.56860), to objective 5, social studies skills (M=0.56758), no statistically significant difference was found (p-value of 0.9680). The pairwise comparison shows that the means of objectives 3 and 5 are the highest (0.56860 and 0.56758 respectively), and the lowest mean of all five was that of objective 1 (history) with a mean of 0.48778.

In conclusion, the results of RQ1 indicate that the null hypothesis of the means of all five objectives are equal, was rejected. Therefore, the tests conducted to address the first research question show that ESL students perform differently on the five test objectives in the social studies TAKS exam.

Research Question 2: Systematic Analysis of Performance by Objective

To answer the second research question, “Is there a systematic difference in the performance of ESL students who remained in the ESL program across objectives in the social studies TAKS exam?” the researcher first identified the students who remained in the ESL program three years later (in 11th grade). Then the two groups of 8th grade percentage raw scores per objective, and 11th grade percentage raw scores per objective were analyzed using a MANOVA. The percentage raw scores per objective were used because there was an unequal number of questions per objective for 8th and 11th grades.

The assumptions of a MANOVA were checked first as follows. The assumption of linearity was satisfied by plotting the residuals of 8th grade raw scores against those of 11th grade. The residuals were subjected to univariate tests of normality by using the Distributions platform in JMP. Because all variables are linearly related and they are not significantly non-normal, it is safe to assume multivariate normality. A Wilk’s Lambda

test of mean differences (a test robust against violations of assumptions) yielded a low p-value of $<.0001$, indicating that the null hypothesis can be rejected, which implies that there is a relationship between at least one X and at least one Y; this meets yet another assumption. Additionally, the correlation between the dependent variables was not high, as it is reported to equal 0.2376, a relatively a low correlation, which rules out strong multicollinearity (which decreases statistical efficiency). As for homogeneity, because there is a large N and equal cell sizes, lack of homogeneity of covariance matrices has minimal impact. According to (Laca, 2012), “if all treatments have the same number of observations,” and “if the ratio of largest to smallest n is less than 1.5, heterogeneity of variance can be ignored, so do not bother testing for it” (p. 9).

Another assumption, which is the sphericity/compound symmetry, which means that the variances of the response variable must be the same at each time point, and that the correlation between repeated measurements are equal regardless of the time interval between measurements, is tested using Mauchly’s Sphericity Test. Since in the JMP output, the two methods used to adjust for lack of sphericity, Greenhouse-Geisser Epsilon and Huyn-Feldt Epsilon, were both equal to 1, which indicated that sphericity wasn’t a problem. Finally, the independence of observations assumption was satisfied per the nature of the data used as each subject’s scores were not influenced or related to the scores of other subjects in each level; each student took the test only once and answered the questions of each objective only once in the spring of each year independently.

As a final note, according to the Mahalanobis Distances analysis which is used to identify outliers, there was a large number of outliers. Too many outliers would affect the results of a MANOVA, however, the researcher deleted all the outliers and performed a MANOVA again (as suggested by Tabachnick & Fidell, 2013), and the same outcome was found. Therefore, the researcher decided to perform the analysis with the outliers to maintain the consistent and equal number of cells per objective given that the results were not altered without the outliers.

The JMP output of the MANOVA shows that the 11th grade students have a higher mean percentage of correct responses than those of the 8th grade students, with the overall mean of 8th graders being 0.479407 and that of 11th graders 0.613399

Table 13 provides the means of percentages of correct responses for the comparison of 8th grade students and 11th grade students in terms of percentage of correct responses by objective. The 11th grade students performed higher in every objective of the five. Figure 5 provides a graph of the means of percentages of correct responses for the comparison of 8th grade students to students in 11th grade on the five objectives of the social studies TAKS test.

Table 13

Descriptive Statistics for the 8th and 11th Grade Social Studies TAKS Test Percentages of Correct Responses per Objective.

	8 th Grade	11 th Grade
<i>Overall mean percent of correct responses</i> (N=24365 for all five objectives)	0.473407	0.613399
<i>Objectives</i>		
Objective 1	0.419912	0.556378
Objective 2	0.482352	0.640171
Objective 3	0.525435	0.622314
Objective 4	0.447569	0.598263
Objective 5	0.491766	0.649870

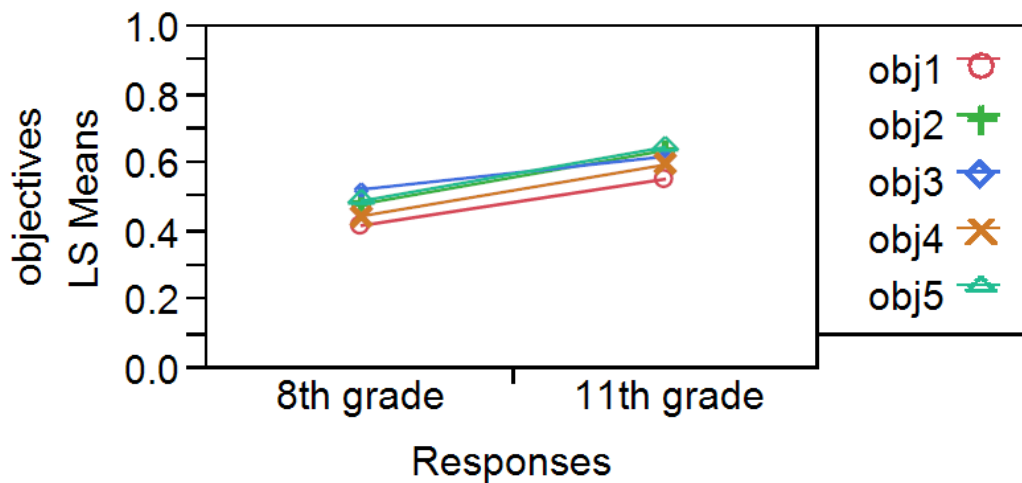


Figure 5. Comparison Between 8th and 11th Grade Social Studies TAKS Test Percentages of Correct Responses per Objective.

The results of the MANOVA show that both the time effect and the objective effect are significant at the alpha=0.05 level (both p-values <.0001). Therefore, there are

significant differences between 8th grade and 11th grade performance (time effect). The two groups differ at any time point, meaning there is an objective effect. There is also a TIME*OBJECTIVE interaction effect (the two groups differ in their responses over time). In addition, the one-way MANOVA revealed a significant multivariate main effect for objective, Wilks' $\lambda = .954323$, $F(8, 48718) = 144.0344$, $p < .0001$.

In conclusion, the null hypothesis of equal means of the percentages of correct responses for all five objectives between 8th and 11th grades was rejected. The results show that ESL student performance, for students who remained in the ESL program, on each of the five objectives, measured by the percent correct responses in the social studies TAKS test, was statistically different. The total number of ESL students who passed the social studies TAKS exam in 8th grade by meeting the state standards, for the years of interest to this study, was 2,351, whereas this number increased to 3,676 in 11th grade. More ESL students performed better as they matured and moved to 11th grade with an increased percentage of more than 56%.

Research Question 3: Analysis of Performance by Question Format

Before addressing the third research question, “Is there a significant difference in the performance of ESL students across item formats which include reading texts (excerpt-based, and bullet-point format) in the social studies TAKS?” the researcher identified excerpt-based and bullet-point question items in the social studies TAKS tests administered in the years 2006 and 2009. For each question (whether excerpt-based or bullet-point) the percentage of correct responses was calculated. A directional

independent t-test was performed to analyze the percentage correct for each question type in the test.

Both the continuous assumption of the dependent variable and the independence of observations assumption in the t-test were met per the nature of the data. First, the dependent variable, which was the percentage correct for each question type is a numerical continuous variable. Second, each observation of the dependent variable was independent of the other observations of the dependent variable since each question item was independent than the others and was not repeated in the same exam.

The normality assumption was met as the Shapiro-Wilk W Test's p-value of 0.1441 indicated inability to reject the null hypothesis of H_0 : the data is from the Normal distribution. However, homogeneity of variance was not the same for the two groups according to Levene's Test (p-value = 0.0259). Finally, no outliers were detected.

Using a significance level of 0.05, the results of the two sided t-test show that there is a significant difference between the two question types (p-value 0.0171). And the one sided directional t-test indicates that the bullet-point question type has a higher mean than the excerpt-based question type (p-value 0.0086).

In conclusion, the null hypothesis of equal means of the percentages of correct responses in the excerpt-based and the bullet-point questions was rejected. Therefore, there was a statistically significant difference in the performance of ESL students across the two question types of interest to this study (excerpt-based and bullet-point questions). Furthermore, the results of the directional t-test indicated that ESL students

chose more correct responses to bullet-point questions than they did for excerpt-based question items.

Research Question 4: Systematic Analysis of Performance on Excerpt-Based Question Type

In order to address this research question, “Is there a systematic difference in the performance of ESL students who remained in the ESL program on test items that are excerpt-based?” the researcher calculated the percentages of correct response in excerpt-based questions for each individual student, who remained in the ESL program, in both the 8th grade and the 11th grade. Descriptive statistics show that the mean percentage of correct responses (N=4873) for excerpt-based questions in 8th grade is 0.4885, and for 11th grade 0.53183. The mean difference is reported to equal 0.04333 with a standard error of 0.00469.

Because of the presence of a number of outliers and the inability to satisfy the normality assumption, a paired t-test (which was initially intended) was replaced by the Wilcoxon Signed Rank non-parametric test for the analysis.

The Wilcoxon Signed Rank test indicated that there is a statistically significant difference between the performance of ESL students who remained in the ESL program on test items that are excerpt-based between 8th and 11th grades; the p-value of <.0001 indicated rejecting the null of equal means. The directional test showed that the same students who took the exam in 11th grade performed better on excerpt-based questions, than when they were in 8th grade (p-value <.0001).

Research Question 5: Analysis of Performance by Demographic Effects

The final research question aimed at finding whether three demographic variables affect ESL student performance on the social studies TAKS exam. The fifth research question “Is there a significant difference in the performance of ESL students in the social studies TAKS exam when specific demographic variables are used as independent variables?” focused on gender, ethnicity, and SES. The dependent variable in this research question was the raw score of students on their social studies TAKS exam. The raw score is a very basic measurement that includes the total number of correct questions the student answers. However, since the number of questions on the social studies TAKS exam varies by grade level, the raw scores were converted to percentages.

To address the final research question including its three sub-questions, the researcher first calculated correlation coefficients to evaluate the correlation between the three independent variables. There seemed to be minimal negative correlation between gender and ethnicity (-0.0146) and between gender and SES (-0.224). Small negative correlation was also found between ethnicity and gender (-0.0146). However, small, yet positive correlation existed between ethnicity and SES (0.2018). The researcher used multiple regression to address the final research question, as it was mainly concerned with exploring the relationship between a dependent variable and a number of independent variables or predictors (StatSoft, Inc., 2012). The three predictors of interest to this research question were gender, ethnicity, and SES. Although the normality assumption was not quiet met, as the Goodness of Fit test (KSL Test) of the residuals

resulted in a probability value of <0.01 , due to the large number of data and the robustness of the F-test to violations of this assumption, this was not of major concern to the results of the multiple regression especially that no outliers were detected (StatSoft, Inc., 2012). Also, as mentioned above, no multicollinearity was found among the three independent variables.

The linear combination of independent variables was significantly related to %Raw score (the dependent variable), as $F(5, 40562) = 190.0456, p < .0001$. All three variables were found to have a statistically significant effect on the dependent variable ($p < .0001$ for all three in the Effects Test in multiple regression). However, only about 2% ($R^2 = 0.02289$) of the variance was accounted for in the dependent variable by the three independent variables. Indeed, R^2 is the percent of the total variation that can be explained by the regression equation, yet, a larger (more desirable) value of R^2 would occur when all of the predictors are included. In this case, only three of the demographic data were considered, consequently the low R^2 value. The $\text{adj}R^2$ value was very low (0.02277) as well as there was a significant lack of fit (p-value of $<.0001$), which says that the regressors or independent variables, don't really explain the variation in the dependent variable very well (Lea, 1997). However, the three independent variables, taken together, were significantly associated with the dependent variable (%Raw score) as the F-Ratio is 190.0456 (with a p-value of $<.0001$). With a p-value of $<.0001$ for all three independent variables (gender, ethnicity, and SES status), the relationships of each regressor with the dependent variable (%Raw score) were statistically significant.

The beta weights of the independent variables, including the dummy variables created for the categorical variable of ethnicity, presented in Table 14, indicate the “expected increase or decrease in the dependent variable, in standard deviation units, given a one standard-deviation increase in independent variable with all other independent variables held constant” (Nathans, Oswald, & Nimon, 2012, p. 3). Beta weights rather than B weights are preferably used for comparison because they are scaled in the same standardized metric (Nathans et al., 2012). Moreover, the beta weights presented in Table 14 are an indication of the total effect of each independent variable and its contribution to a multiple regression equation; basically, it shows the importance of each variable (Nathans et al., 2012). The table shows that the independent variable of ethnicity, the Asian ethnic classification in specific, with a beta weight value of 0.09991 has the most effect on the dependent variable, followed by SES, and finally gender. Nevertheless, the beta weight value may not fully explain the contribution of independent variables, as the variance may be shared by different independent variables (Nathans et al., 2012).

Table 14

Simultaneous Multiple Regression Analysis Summary for Gender, Ethnicity, and SES Predicting % Raw Score in the 2003, 2006, and 2009 Social Studies TAKS Tests.

Source	B	SE	Beta
Gender	0.02222	0.00177	0.06182
SES	-0.03412	0.00265	-0.06536
Native American	0.00000	0.00000	0.00000
Asian	0.08936	0.00452	0.09991
African American	0.04860	0.01063	0.02245
White	0.05591	0.01037	0.02660

Note. $R^2 = .02289$; $F(5, 40562) = 190.0456$, $p < .0001$.

The Beta weights in Table 14 suggest that when other variables are held constant, the gender coefficient of 0.02222 shows us that males' %Raw scores average a 0.02222 times higher than females' %Raw scores. Also, with the dichotomous variable of SES status (advantaged vs. disadvantaged), the SES coefficient of -0.0375, with advantaged coded 0 and disadvantaged coded 1, means that with all other variables held constant, the %Raw scores of disadvantaged ESL students average 0.03412 times lower than the advantaged ESL students. As for ethnicity, since the Hispanic group was the control group (as a dummy variable was not created for it), the beta weights show that all the other ethnicities performed better than this group. The Native American ethnic group did not include any participants. However, the Asian ethnic group, with all other variables held constant, performed 0.09991 times better than the Hispanic group. The African American ethnic group also outperformed the Hispanic ethnic group by 0.02245 times. Finally, the %Raw scores of White/ non-Hispanic ESL students were 0.02660 times higher than the ESL Hispanic ethnic group.

In conclusion, the results of the multiple regression reject the null hypothesis of equal %Raw scores in the social studies TAKS exam for ESL students of different genders, varying ethnicities, and different SES. Thus, there is a statistically significant difference in the performance of ESL students in the social studies TAKS exam when gender, ethnicity, and SES are used as independent variables.

Summary of Findings

The several statistical tests used to address the research questions of interest in this study show that there was a mean difference in the performance of ESL students in the five objectives of the social studies TAKS exam. The mean difference of objectives 3 (economic and social influences) and objective 5 (social studies skills) were not statistically significant, whereas the other three objective means were. ESL students in 8th and 11th grade in the years 2003, 2006 and 2009 performed best on objectives 3 and 5, and poorest on objective 1 (history). The test results also revealed that there was a systematic difference in the performance of ESL students who remained in the ESL program across objectives in the social studies TAKS exam. When comparing the performance of 8th grade ESL students to their performance in 11th grade 3 years later, while still participating in the ESL program, a time effect, an objective effect and an interaction effect were found.

In terms of question item types, the results of the statistical tests showed that the bullet-point question type yielded a higher % correct mean than the excerpt-based question type, as ESL students chose correct answers for bullet-point questions more than they did for excerpt-based questions. However, specifically for excerpt-based

questions, there was a systematic difference that was presented through statistical tests, where ESL students in the eleventh grade performed better on excerpt-based questions than they did three years earlier in 8th grade.

Finally, in predicting the percentage of raw scores in the social studies TAKS exam, the researcher found that the independent variables of gender, ethnicity, and SES, were statistically significantly related to the dependent variable (%Raw score).

CHAPTER V

SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS

The importance of the subject content of social studies has been explored and realized by many in the realm of education and among the general public. Since the State of Texas includes it as one of the major subjects tested on its state TAKS exams, the full exploration of the subject is important to help teachers address the needs of ESL students in order to support their success. Because the social studies TAKS test is divided into five objectives, the performance of ESL students on these objectives was analyzed. In addition, question types that include reading text, thus seemingly more use of the English language, may have been challenging to ESL students. Therefore, the performance of ESL students on these question types was examined as well.

This chapter provides a summary of the purpose of the study and a restatement of the research questions. A discussion of the results, organized by research question, follows with emphasis on conclusions. Finally, the last section of this chapter is dedicated for general implications for both theory and practice, ending with a discussion of the limitations of the study and future recommendations.

Purpose of the Study

The purpose the study was to analyze a number of questions related to ESL students' performance on social studies TAKS exams with the intention of finding points of weaknesses to be addressed by content-area teachers or ESL teachers. Both content (reflected in the objectives) and question item format were explored to determine if there

is a relationship between ESL status and the students' ability to successfully pass the social studies TAKS exam. The study sought to determine if the performance of ESL students was influenced by the organization of a question that included some form of enboxed reading text, given that ESL students are not fully proficient in the language yet, as implied by their placement in the ESL program. It was also the intention of this study to examine performance across the years, while students remained participants in the ESL program, both by social studies objective and by question type, in order to determine if a pattern exists in performance as a function of maturation of the ESL student. A final purpose of the study included exploring the influence of demographic effects (gender, ethnic classification, and SES) on the overall performance resembled by the total raw score of ESL students in the social studies TAKS exam.

Conclusions and Discussion of Research Questions 1& 2: Analysis by Objective

The first goal of this study was to investigate the difference in the performance of ESL students on the social studies TAKS exam based on content, which is represented by objectives. The first two research questions were directed to that goal, although to different extents.

In research question one, since the homogeneity assumption was not met as discussed in the results chapter, the Welsh test was used instead of an ANOVA. The test results showed a statistically significant difference at the $p < .05$ level in the percentage of correct answers for all the objectives except for objectives 3 (economics) and objective 5 (social studies skills). This finding is not really explainable as the two objectives are not related in content and expectations (TEA, 2007b).

Descriptive statistics of research question one show that objectives 5 (social studies skills) and objective 3 (economics) have the highest mean of percentage of correct answers for ESL students in grades 8 and 11 in the years 2003, 2006, and 2009. This is in accordance with Smith's (2012) analysis that attests to objective 5 standing out as a unique objective and different from the other four that emphasize accumulation of knowledge. The TAKS Study Guide for grade 11 (2007) states that in objective 5 students are expected to use critical-thinking skills. Moreover, the mere nature of objective 5 being an incorporating unit for skills from different content areas such as mathematics and science, which does not limit its scope to the social studies subject matter (Smith, 2012), makes it more feasible for ESL students. Thus, in alignment with Smith's (2012) analysis, and the identification of student expectations and sample topics provided by the TAKS Study Guide (2007), the results related to objective 5 are explained. On the contrary, the second objective to top the list as having the highest performance rates (which is in fact not statistically significantly different from the performance of ESL students in objective 5), has not been addressed by previous research. According to the TAKS social studies Study Guide for grade 11 (2007), objective 3 deals with economic and social influences on historical issues and events. Although this objective is closely related to objective 1 (U.S. history), which was the least favorable among ESL students in terms of performance, students still highly performed on that objective.

Objective 1 in which students are expected to show understanding of issues and events in U.S. history (TEA, 2007b), was the most challenging to ESL students in 8th

and 11th grades according to the descriptive statistics of research question one as it had the lowest mean percentage of correct answers. Part of the challenge lies in the vocabulary used to convey the information related to this objective as suggested by previous research (e.g. Chamot & O'Malley, 1994; Cho & Reich, 2008; Gonzales, 1988). Another part is a result of culture or cultural differences as discussed by many (Cho & Reich, 2008; Duff, 2001; King et al., 1987; Reilly, 1988; Smith, 2012). Finally, it could be part due to the idea of history being the result of information accumulation throughout the years when ESL students were not enrolled in Texas schools (Cho & Reich, 2008; Duff, 2001; Smith, 2012; TEA, 2004a). Therefore, in accordance with the general challenges discussed, historical events specific to the U.S. could be the most challenging to ESL students, as the results of RQ1 suggest.

However, despite the fact that objective 3 is related to historical issues and events in the U.S. specifically, students chose more correct responses in objective 3. High performance was detected in the study although memorization, rather than critical thinking skills as is the case with objective 5, is emphasized in this objective as more than half of the verbs used to describe it fall in the lower half of Bloom's Taxonomy according to Smith (2012). The high performance in objective 3 is also contrary to the fact that social studies vocabulary comes from abstract ideas from areas such as economics which increases the difficulty of it (Gonzales, 1988). Previous research has not focused on the economic and social influences part of the social studies TAKS, which leaves this finding unexplainable, and opens venues for future research.

Dissimilar to objective 1, in which knowledge is specific to the U.S. and is generally an accumulation of information, Cole (2007) argues that geographic learning (which is the emphasis of objective 2) starts even before a student is part of an American classroom. The results of RQ1 show that the performance of ESL students on objective 2 (geography) was relatively high, following objectives 3 and 5. This may be due to Cole's (2007) explanation of the universality of spatial competence which is a skill that starts developing at an early age, even before being part of any school system. Therefore, skills needed to be able to answer objective 2 related questions may not be necessarily a result of the schooling system in the U. S.; thus, ESL students not being part of the American educational system at an early age, does not negatively influence their performance on geography questions. Nevertheless, the eighth grade Social Studies Booklet (TEA, 2004a) indicates that the study of geography and history are strongly related as they explain the where and the why of events.

A second aim of this study was to explore the possible effect of maturation of ESL students who remain in the ESL program on their performance on the social studies TAKS exam. A systematic explanation was sought in that regards by matching students' performance in 8th grade with their performance in 11th grade by objective. The results indicated that ESL students performed statistically significantly better in 11th grade after at least five to six years of being in the American schooling system (three years being exempt from taking the TAKS, and three years since their test taking in 8th grade). This finding supports previous research regarding English language acquisition suggested by Cummins (1981), as he stated that five to seven years are required to gain CALP in the

second language. It also supports Collier and Thomas' (1989) assertion of the gradualism of the process of language proficiency in L2.

Conclusions and Discussion of Research Questions 3 & 4: Analysis by Question Type

This study also aimed at comparing question item types and their influence on performance on the social studies TAKS exam. The two question items chosen for analysis in this study were two types that include a fair amount of reading text which is confined in a text box. These two types were specifically chosen due to the linguistic challenge ESL students face in terms of reading texts because of their limited English proficiency (Bernhardt & Kamil, 1995; Cho & Reich, 2008; Coltrane, 2002; Duff, 2001; Eslami & Plett, 2008; Gonzales, 1988; Shin, 2002). The aspect selected for exploration in these two question types was the organization of the text incorporated in them.

The text that is included inside the boxes, as part of the question text, was analyzed according to the way it was organized, whether in bullet points or in an excerpt format. The first question item, bullet-point based, included an enumeration typographical signaling device which was the bullet points, whereas the other type (excerpt-based) did not. The results of this part of the study (research question 3) which indicated better performance on bullet-point based questions, aligns with previous research results in which texts that utilized an enumeration device yielded better results; that is because the enumeration device facilitated recall (Lorch & Chen, 1986), and because it drew attention to signaled information (Narveson, 2001). In addition, although one variable cannot be the sole or major influence on students' performance in reading

texts, Narveson's (2001) findings showed that these cues help weaker readers – in this case, ESL students. Therefore, the findings of this study are in accordance with Narveson's (2001) results.

As for excerpt-based questions, student performance was lower than the performance on bullet-point questions most likely due to the nature of this type of questions. Since excerpts are shorter versions of longer passages from some kind of source, which is usually a book, a speech, or a government document, they include a considerable amount of reading text that is lumped in a chunk rather than organized in a more visually appealing manner such as the case with bullet-point questions. Moreover, the fact that these texts come from a primary authentic source which is written for the general adult public such as the Declaration of Independence, may cause the increased difficulty level in these question types. This falls in line with Das' (2007) conclusion of his study relating ELL reading scores and their mathematics scores for grades 3 through 11. Das (2007) found a strong positive correlation ($r=0.852$) between ELLs' performance in reading and in math; meaning, if a student performed poorly in reading, they most likely performed poorly in math. That may be the result of the amount of reading present in other content-area exams such as mathematics, science, or social studies. Furthermore, as students nowadays are exposed to more technology in the classroom, PowerPoint presentations are widely used by teachers. PowerPoint presentations are usually in the format of bullet points rather than in excerpt-looking text. Thus, ESL students are more likely to find the bullet-point format more familiar and easier to follow and respond to.

A more detailed observation of the type of questions ELLs found more challenging may further support the current research, specifically on questions with relatively more text in them. Moreover, a brief look at the objectives these excerpt questions fall under, shows that most of the excerpt questions lie under objective 5 (35% of the questions), followed by objectives 1 and 4 equally (23%), whereas none of the excerpt questions were under objective 2 (0%). Further research may add to the relationship between excerpt-based questions and the objectives they lie within in order to assist social studies teachers in addressing these types of questions in the classroom to possibly enhance performance.

Nevertheless, looking at the brighter side of excerpt-based questions, the results of research question four indicate that there is a pattern in the performance of ESL students who remained in the ESL program on excerpt-based questions through the three years between 8th and 11th grade. Their performance in 11th grade is statistically better than their performance in 8th grade, which is in line with the results of research question two discussed earlier, as ESL students' performance by objective improved as a result of time and maturation. In terms of excerpt-based questions that require a significant degree of language proficiency, the language learning gained from being in the schooling system in the U.S. for a number of years (at least six years as explained earlier) has most likely helped ESL students in performing better on excerpt-based questions as they have developed more of the CALP (Collier & Thomas, 1989; Cummins, 1981).

The conclusions from these two research questions showed that ESL students performed differently on the two question types of interest to this research, favoring the

bullet-point questions over the excerpt-based questions. The nature of excerpt-based questions in incorporating text written for the general English-speaking public, may have contributed to the challenge ESL students found in answering them correctly.

Nevertheless, the maturation of ESL students was evident in their improved performance after a three-year period between 8th and 11th grades.

Conclusions and Discussion of Research Question 5: Analysis of Demographic Effects

Finally, this research examined three different independent variables and whether they affect the performance of ESL students on the social studies TAKS. Gender, ethnic classification, and social economic status were studied to find if they had an effect on the dependent variable of percentage raw score. The multiple regression analysis indicated that the three independent variables had a statistically significant effect on the dependent variable (%Raw score) ($p < .0001$ for all three).

The independent variable of gender had the least beta weight value (0.062686), meaning it had the least effect of all three on the dependent variable. However, one issue with beta weights is that they fail to recognize the possible associations between the independent variables (in this case gender, ethnicity, and SES) (Nathans et al., 2012). The ‘shared variance’ as Nathans et al. (2012) calls it, does not totally capture the contribution of the independent variables. This is also shown in previous research such as that of Silva (2010) who recognized the total influence of the interaction between gender, SES, and limited English proficiency on student performance.

The gender coefficient showed that ESL male students scored higher than their female counterparts in the social studies TAKS test. The findings of this part of this study further support previous research in the field specific to social studies (e.g. Silva, 2010; Quezada, 2008), yet, contrary to the findings of other scholars in other areas of study such as language arts and mathematics where females outperformed males in general (e.g. Chudowsky & Chudowsky, 2010; Ikegulu, 2004; Quezada, 2008; Reiss, 2005). In addition, the findings of this research are in contrast to the statewide statistics reported by the TEA in the Academic Excellence Indicator system reports for the years of interest (*AEIS 2002-2003; AEIS 2005- 2006; AEIS 2008- 2009*). These reports show no difference in gender for the social studies assessment in 2009, and for 8th grade in 2006. A more stark contrast is in the TEA reporting of higher female performance for the year of 2003 as shown in Table 15.

Table 15
TAKS Passing Rates on the Social Studies TAKS Test by Gender for the Years 2003, 2006, and 2009.

	2003		2006		2009	
	<i>8th Grade</i>	<i>11th Grade</i>	<i>8th Grade</i>	<i>11th Grade</i>	<i>8th Grade</i>	<i>11th Grade</i>
Male	92.6%	89.1%	84%	95%	92%	97%
Female	93.7%	91.2%	84%	93%	92%	97%

(TEA, 2003a; TEA, 2006a; TEA, 2009a)

Indeed, the NCLB assessment represented by standardized tests such as the TAKS test is aimed at closing the achievement gap between student groups. Gender gaps

have a more popular audience in the language arts and mathematics content areas, but are also worth recognizing across other areas such as social studies. Information on gender gaps could give educators an insight on the issues their students are dealing with in different areas of study, consequently leading to enhanced educational opportunities and materials that are suitable for both genders to help both succeed.

Another finding from this research question deals with the independent variable of socio-economic status. The researcher used the definition provided by the TEA and its categorization of students into the two groups of *economically advantaged* and *economically disadvantaged*. The findings indicated that the economically disadvantaged group of ESL students performed lower than the economically advantaged ESL student group. This finding is in accordance with the findings and conclusions of many, if not all, the research performed in the field of education (e.g. Aguilar, 2010; Das, 2007; Heier, 2011; Ikegulu, 2004; Jaska et al., 2009; Langen et al., 2006; Nichols, 2003; Quezada, 2008; Silva, 2010; Sirin, 2005; Slaughter, 2007; Stewart, 2009; White, 1982; Zamarripa, 2009). More specifically, this study supports the findings of Quezada (2008) and Silva (2010) in relation to the content area of social studies; both studies found that students with lower SES performed poorer.

Moreover, the results of this part of the study, similar to Das' (2007) opinion, do not consider a student's SES to be a primary reason for the low performance of ELLs on high-achievement tests. That is, because the effect is not overwhelmingly great (shown by the very low $\text{adj}R^2$ value) and as evident in its beta weight value which is lower than that of ethnicity.

SES has even greater influence according to Jaska et al. (2009) as the percentage of economically disadvantaged students in a school influences students' performance on TAKS exams negatively. This may challenge educators in finding ways to minimize the gap between SES student groups as these economic influences are out of their control. Consequences of such a statement may affect teachers and the belief in their inability to influence student achievement with their teaching efforts (meaning that student success is beyond their control) (Chadwick & Harris, 2009). However, the curriculum and pedagogy, in general, do not make use of lower SES students' existing knowledge and strengths. The schools may be blamed for such poor performance of low SES students as schools and teachers often treat this student segment differently, and may not challenge them enough through the curriculum and the instruction they offer.

Nevertheless, students from low-income families do undergo circumstances that are beyond a teacher's control. They perform lower than higher SES students due to family pressure to work, for example, not having a role model or support from an adult, and possibly a lack of educational resources at home causing limited background knowledge in general and in vocabulary specifically. In regards to social studies in particular, the inadequate background knowledge low-SES students have due to the limited access to vocabulary and educational opportunities may be a cause of their lower achievement in standardized tests, as discussed by Aguilar (2010).

Finally, ethnicity was the last independent variable to be examined in this study. According to the beta weights of all three independent variables studied in this research question, ethnicity had the highest value, meaning it had the biggest effect, of all three,

on the dependent variable (%Raw score). This conclusion falls along the same lines as those found by Quezada (2008) where ethnicity was the strongest predictor of achievement in the content area of social studies with the Hispanic group performing the lowest. In contrast, Nichols' (2003) study showed that other factors such as economic status and poor attendance, as well as failure in earlier school years, were more important factors than ethnic background in explaining low performance. Nevertheless, the interrelationship between demographic characteristics and performance helps better understand their effect in order to enhance educational environments for all the students in our schools. In relation to that, minority ethnic groups tend to be categorized as being part of lower socio-economic environments (Fry & Gonzales, 2008), which are also negatively related to achievement as shown earlier.

The results of the multiple regression analysis indicated that ethnicity has an effect on performance (measured by %Raw score). This is in accordance with Quezada's (2008) conclusions for the content area of social studies, as well as in other content areas such as mathematics as concluded by Roane (2008); both Roane's (2008) study and the NCTM (2000) reports show the existence of achievement gaps based on ethnicity in the content area of mathematics. Specific to ESL students, the still developing English language proficiency these students and their families possess is a major influence on their achievement (Altshuler & Schmautz, 2006).

In summary, the final research question indicated a significant independent factor effect of all three demographic variables (gender, SES, and ethnicity) on the

performance of ESL students on the social studies TAKS test (measured by %Raw score).

General Implications for Theory and Practice

As is the case with any subject matter, teachers are the student advocates when it comes to standardized tests. Classroom teachers work on improving students' performance in general and more specifically focusing on areas that are challenging to the students. Teachers who have ESL students in their classrooms are under yet more pressure to help their students succeed in such exams, especially the high-stakes standardized ones that are mostly state-wide exams due to the expectations and requirements of the NCLB (Bouroughs et al., 2005; Chargois, 2008; Cizek & Burg, 2006; Coltrane, 2002; Funkhouser, 1990; Ikegulu, 2004; U.S. Department of Education, 2002).

Although the five test objectives of the social studies TAKS exam do not offer a complete explanation of student performance (as per the R^2 value shown in the results of RQ1), they are still an indication of student performance and what areas students find more challenging. The results of this study show statistically significant difference in the performance of ESL students by objective on the test, with higher performance on certain objectives and lower performance on others. These results are significant for classroom teachers of ESL students in considering which areas and subjects to focus on in the classroom and which objectives to emphasize throughout the school year in order to better prepare the students for the test. The idea is not to teach to the test, although it may sound like it is, but rather to assist ESL students in areas they need more help with.

History, for example, which is represented by objective 1 in the social studies TAKS, is shown to be most difficult for ESL students. Many factors discussed earlier may contribute to that (e.g. Chamot & O'Malley, 1994; Cho & Reich, 2008; Duff, 2001; Gonzales, 1988; King et al., 1987; Reilly, 1988; Smith, 2012). One of the major reasons, as discussed earlier, is the difficulty of concepts in History, which teachers should attend to more. However, regardless of the reasons, this is an issue that needs to be addressed by teachers and social studies departments in schools. More support is needed for ESL students in the area of history to help them improve their performance on the social studies TAKS exam or any replacement of it such as the STAAR. This does not mean that other objectives are to be ignored or to be given less attention, it is just to inform educators of what this student segment finds more challenging in order to address it in the most suitable manner depending on the students, the teachers, and the resources available. One way ESL students might find history more accessible is if it included more world history, to which ESL students may have already had exposure through their home country's curriculum.

As for question type, educators are encouraged to expose students to various excerpts as part of classroom activities and to place extra emphasis on them as they may be foreign to some students in both content and language. In terms of boxed question items with reading text, excerpt-based questions were more challenging to ESL students in comparison to bullet-point questions. The study found significant differences between excerpt-based and bullet-point question types based on students' correct answers for each type of questions in the social studies TAKS tests for the years 2006 and 2009.

Thus, the study indicated that ESL students were less successful on excerpt-based questions than they were on bullet-point questions. Although the performance of this student group increased over the years from 8th grade to 11th grade, it still created some kind of an obstacle to ESL students, thus, qualifying it as an issue teachers must look into and rethink the way in which it is addressed and taught in class.

While excerpt-based questions do not necessarily solely explain the overall performance of students on the social studies TAKS exam, the fact that they exist in a varying density in each test of every year is significant for educators to consider when teaching social studies to ESL students. It is essential that mainstream teachers are knowledgeable of and trained to enrich ESL students' abilities in developing higher-level cognitive-academic skills in order to facilitate their transition from the ESL classes into the mainstream classes (Collier & Thomas, 1989).

This research further supports the theories of Cummins (1981) and Collier and Thomas (1989) in language acquisition related to CALP. Educators in decision-making positions in the government and on the state level ought to look at studies such as the present one that disqualify the three year period of the exemption rule to be sufficient for students who speak English as a second language in standardized state examinations. An extended period of time that goes beyond three years gives ESL students more time to acquire the language necessary to succeed in the educational system in the state in which they live. The systematic evidence in this study for the objective level, question-type level (excerpt-based specifically), and overall performance level support that. ESL students who were classified as ESL students and remained in the ESL program for over

three years (about six to be more accurate, as explained before), showed better overall performance in each objective of the five, thus, increased performance overall on the social studies TAKS test, in addition to better performance in excerpt-based questions.

Other implications are related to social studies instruction. This study recommends more exposure to excerpts and documents found in our everyday lives or in topics related to the social studies curriculum used. That is first, to increase linguistic learning as the language used may incorporate more sophisticated language forms and vocabulary, and second, to familiarize ESL students as well as non-ESL students with the nature of such documents and the language used in them. Finally, more exposure to excerpts and excerpt-based questions gives students the opportunity to develop higher cognitive skills, as the analysis of such authentic texts requires more sophisticated critical-thinking skills. Such skills help create students who are critical thinkers who grow into responsible educated citizens in our society; this in turn is stated as one of the important goals of the study of social studies as suggested by Burroughs et al. (2005).

As for instruction in the social studies content area, and in regards to ESL students, teachers are encouraged to use social studies skills (objective 5) to address the other four objectives. Teachers may find this approach to be helpful for mainstream students as well. Since ESL students perform best on the only objective that incorporates skills and multiple disciplines, this should be part of the teaching strategy of teachers. Knowledge-based instruction can be infused with the skills-based instruction to provide the students with positive learning experiences, which ultimately lead to better performance on tests.

King et al. (1987) and Short (1994) offer an extensive list of ways to infuse skills with content in the area of social studies. Skills including outlining, note-taking, summarizing (Short, 1994), paraphrasing, webbing to illustrate relationships, drawing diagrams to show cause-and-effect relationships, interpreting cartoons, and reading maps (King et al., 1987) can be integrated in the social studies classroom to understand content. Moreover, skills such as describing and analyzing pictures and visual aids (King et al., 1987), along with other skills such as dictionary practice, library search, and identifying definitions through context (Short, 1994), can assist in explicit vocabulary instruction to simplify content.

Content areas such as language arts and social studies share skills that may be developed in both classes. According to King et al. (1987) and Short (1994), in language arts classes, students learn how to order events to show sequence, which is similar to the study of historical events through storytelling skills for example. Another example is in relation to geography as the students learn to describe the characters and the setting in both language arts and in social studies. In Short's (1994) study, participating social studies teachers created activities such as comparing and contrasting two historical characters, and at times interpreted and analyzed poems in their social studies classes to facilitate understanding. Along the same lines, by integrating the writing process skills that include brainstorming, organizing, and revising, ESL students may be able to access the content of the social studies class in a more effective way (King et al., 1987; Short, 1994). Furthermore, higher-level thinking skills may be infused in the content through

teaching skills such as predicting, making assumptions, reading, and evaluating through classroom activities, as suggested by King et al. (1987).

Finally, mainstream teachers in general, including social studies teachers, are to be sensitive towards the needs of ESL students in general and to their abilities in the content area of social studies. That is, to better serve this student segment and to provide them with learning experiences that are related to their knowledge and lives.

With the reauthorization of the NCLB act, this study recommends that there be less accountability for ELLs' achievement in general due to the linguistic and prior knowledge challenge they face. ELLs are to be held accountable, but studies such as the current one support the need of giving them more time before they are. Although the content and language used in social studies tests are a burden on ESL students, the benefits of social studies (as have been discussed earlier) show that it is a content area of importance in creating a responsible citizen. Therefore, this study recommends that social studies be tested for accountability to help develop students that are well-rounded academically and socially, instead of detaching them from our society.

Discussion of Limitations and Future Recommendations

A limitation of this study resulted from the exploratory nature of the study based on the existing information available of students participating in an ESL program in the State of Texas who were given the social studies TAKS examination. The researcher had no control over the academic ability of the students both in the English language and in the content areas, instructional design of the program they were part of, teachers in the programs, or any other factor that may have led to or prohibited academic achievement

such as age of arrival to the U.S. These individual differences may be of greater impact on performance, but were not available to the researcher. In addition, school information specific to the type of program, quality of the program and the teachers, and overall performance of the general population in the schools were not discussed in this study as the information provided was limited to not include such information. This information may be provided by the TEA if requested, based on certain criteria set by the TEA, thus, the possibility for future research to examine the role of the school these ESL students study at on their performance in social studies standardized tests. The researcher chose not to include the school information because the current research aimed at assisting ESL students in general regardless of the status of their schools. If the school information was to be provided, other information that was of interest to this research may have not been available to the researcher according to the TEA (Jennifer J. Eaton, personal communication, June 30, 2011). Thus, the researcher chose to limit the study not to include detailed school information.

In addition, grade choice (8th and 11th grades) was a delimitation of this study due to the importance of these grades in a student's school years. Future research may focus on the psychological status of ESL students and the pressure of passing the social studies portion of their TAKS tests. That could be attained by examining the difference between ESL students' performance in 10th grade versus their performance in 11th grade, as there is no serious consequence of how they perform in 10th grade. Another area to explore is the performance of students who took the Alternative TAKS exams and those who needed any sort of accommodations to take the exam. In addition, the same study may be

repeated for students who are labeled as LEP, despite their participation in an ESL program. Limited English Proficiency status is given to students in and out of the ESL program, which covers a bigger student population than the one chosen for this study.

Another area of future research this study may lay foundation to, focuses on question item analysis in terms of typography and organization. As this study was limited to two types of question items, there is much more to explore in that regards. The effect of typographic cues on ESL versus non-ESL students can be examined as well.

Future studies might expand beyond the social studies TAKS exam and focus on its relationship to the other parts of the TAKS exam. Das (2007) examined the relationship between ELL students' performance in reading and their performance in mathematics and found a strong positive relationship. Such a study may be repeated with other content areas including social studies to observe whether there is a possibility of a trend among ELLs or not in TAKS exams in specific, or any other standardized test.

Qualitative study exploration of the content area of social studies and its portion in the TAKS exams ought to be given more attention in contemporary research. Social studies is not regarded as one of the higher-tier content areas such as reading, mathematics, and science, yet its inclusion in standardized state exams such as the TAKS make it a challenging task for the teachers to teach the subject. An insight to what teachers find challenging in teaching social studies in general, and to an ESL population in specific may be explored qualitatively. Also, qualitative research may affirm the results of the current study in understanding why certain objectives are considered perhaps more challenging or difficult than others. Along the same lines, qualitative

research may help in the exploration of the objectives ESL students found favorable and the reasons why.

Finally, future studies may repeat the procedures of the current study with smaller sample sizes, as the large sample size may have caused the statistically significant results in this study.

As a final observation, since the Texas Education Agency has replaced the TAKS test with the State of Texas Assessments of Academic Readiness (STAAR), a similar analysis of that test may be conducted in all tested subjects, including social studies in eighth grade and at the high school level, to examine the effect of language ability status on test items and objectives as measured by performance.

REFERENCES

- Aguilar, R. M. (2010). *The relationship between Hispanic teachers and Hispanic student academic achievement in Texas*. (Doctoral dissertation). Retrieved from ProQuest Digital Dissertations. (UMI 3405817).
- Altshuler, S., & Schmutz, T. (2006). No Hispanic student left behind: The consequences of “high stakes” testing. *Children & School*, 28 (1), 5-14.
- Bachman, L. F., & Palmer, A. S. (1982). The construct validation of some components of communicative proficiency. *TESOL Quarterly*, 16 (4), 449-465.
- Bedford, M. M. (2007). *Exploring current civic knowledge of 8th grade students in South Texas as measured by the Texas Assessment of Knowledge and Skills*. (Doctoral dissertation). Retrieved from ProQuest Digital Dissertations. (UMI 3322108).
- Bernhardt, E. B., & Kamil, M. L. (1995). Interpreting relationships between L1 and L2 reading: Consolidating the linguistic threshold and the linguistic interdependence hypothesis. *Applied Linguistics*, 16(1), 15 –34.
- Bernhart, S. (1986). Seeing the text. *College Composition and Communication*, 37 (1), 66-78.
- Black, W. R., & Valenzuela, A. (2004). Educational accountability for English language learners in Texas: A retreat from equity. In L. Skrla & J. Scheurich (Eds.), *Educational equity and accountability: Paradigms, policies, and politics* (pp. 211-230). New York, NY: RoutledgeFalmer.
- Britton, B. K., Glynn, S. M., Meyer, B. J. F., & Penland, M. J. (1982). Effects of text structure on use of cognitive capacity during reading. *Journal of Educational Psychology*, 74, 51-61.
- Burroughs, S., Groce, E., & Webeck, M. L. (2005). Social studies education in the age of testing and accountability. *Educational Measurement: Issues and Practice*, 24 (3), 13-20.
- Carrell, P. L. (1984). Schema theory and ESL reading: Classroom implications and applications. *The Modern Language Journal*, 68 (4), 332-343.
- Case, R. (2006). How to assess language in the social studies classroom. *The Social Studies Journal*, 97 (1), 41-48.

- Chadwick, M. & Harris, S. (2009). An exploration of the relationship between the Texas Assessment of Knowledge and Skills (TAKS) math achievement and student resiliency. *Graduate Research Journal, 1*, 1-24. Retrieved from <http://www.grjsite.com/volume-1/19-an-exploration-of-the-relationship-between-the-texas-assessment-of-knowledge-and-skills-taks-math-achievement-and-student-resiliency.html>
- Chamot, A. U., & O'Malley, J. M. (1994). *The CALLA handbook: Implementing the cognitive academic language learning approach*. New York, NY: Longman.
- Chargois, T. (2008). *Student achievement: Identification of impact variables*. (Doctoral dissertation). Retrieved from ProQuest Digital Dissertations. (UMI 3415929).
- Cho, S., & Reich, G. A. (2008). New immigrants, new challenges: High school social studies teachers and English language learner instruction. *The Social Studies, 99* (6), 235-242.
- Chudowsky, N. & Chudowsky, V. (2010). State test score trends through 2007-2008: Are there differences in achievement between boys and girls? Washington, D.C: Center on Education Policy. Retrieved from <http://www.cep-dc.org>.
- Cizek, G., & Burg, S. (2006). *Addressing test anxiety in a high-stakes environment: Strategies for classrooms and schools*. Thousand Oaks, CA: Corwin Press.
- Cole, B. (2007). Spatial competence in Texas high school students. *University Honors Program*. Paper 58. <http://ecommons.txstate.edu/honoprogram/58>
- Collier, V. P., & Thomas, W. P. (1989). How quickly can immigrants become proficient in school English? *The Journal of Educational Issues of Language Minority Students, 5*, 26-38.
- Coltrane, B. (2002). *English language learners and high-stakes tests: An overview of the issues*. Washington, D.C: Center for Applied Linguistics. Retrieved from http://www.cal.org/resources/digest/digest_pdfs/0207coltrane.pdf
- Cummins, J. (1979). Linguistic interdependence and the educational development of bilingual children. *Review of Educational Research, 49*, 222-251.
- Cummins, J. (1981). The role of primary language development in promoting educational success for language minority students. In California State Department of Education (Ed.). *Schooling and language minority students: A theoretical framework* (pp. 3-49). Los Angeles, CA: National Dissemination and Assessment Center.

- Curtin, E. (2005). Teaching practices for ESL students. *Multicultural Education Journal*, 12 (3), 22-27.
- Das, K. P. (2007, March 26). An analysis of mathematics and ELL perspective. *The Hispanic Outlook in Higher Education*, 17, 14-16.
- Diaz, M. d. (1995). *Effect of text layout on students' immediate retention and use of concepts*. (Doctoral dissertation). Retrieved from ProQuest Digital Dissertations. (UMI 9614930).
- Dietz, S. (2010). State high school tests: Exit exams and other assessments. Washington, D.C: Center on Education Policy. Retrieved from <http://www.cep-dc.org>.
- Duff, P. A. (2001). Language, literacy, content, and (Pop) culture: Challenges for ESL students in mainstream courses. *The Canadian Modern Language Review*, 58 (1), 103-132.
- Else-Quest, N. Hyde, J., & Linn, M. (2010). Cross national patterns of gender differences in mathematics: A meta-analysis. *American Psychological Association Bulletin*, 136 (1), 103-127.
- Eslami, Z., & Plett, B. (2008). High-stakes testing and ELLs: L2 reading research applied to the Texas Assessment of Knowledge and Skills. *TABE Journal*, 10 (1), 1-27.
- Flanagan, C., & Faison, N. (2001). Youth civic development: Implications of research for social policy and programs. *Social Policy Report*, 15 (1), 3-14.
- Flores, S. M., Batalova, J., & Fix, M. (2012). The educational trajectories of English Language Learners in Texas. Washington, D.C: Migration Policy Institute.
- Funkhouser, C. (1990). *Education in Texas: Policies, practices, and perspectives* (5th ed.). Scottsdale, AZ: Gorsuch Scarisbrick Publishers.
- Gall, M., Gall, J., & Borg, W. (2003). *Educational research: An introduction*. Boston, MA: Allyn & Bacon.
- Glaser, B. G. (1963). Retreading research materials: The use of secondary analysis by the independent researcher. *The American Behavioral Scientist*. 6 (10), 11-14.
- Glass, G. V. (1976). Primary, secondary, and meta-analysis of research. *Educational Researcher*, 5 (10), 3-8.

- Gonzales, F. (1988). Integrating the ESL student into the content area classroom: Training module II. San Antonio, TX: Intercultural Development Research Association. Retrieved from http://www.eric.ed.gov/ERICWebPortal/search/detailmini.jsp?_nfpb=true&_E_RICExtSearch_SearchValue_0=ED322226&ERICExtSearch_SearchType_0=no&accno=ED322226
- Gwet, K. L. (2008). Computing inter-rater reliability and its variance in the presence of high agreement. *British Journal of Mathematical and Statistical Psychology*, 61, 29-48.
- Hakim, C. (1983). Research based on administrative records. *The Sociological Review*, 31 (3), 489-519.
- Haney, W. (2001). Revisiting the myth of the Texas miracle in education: Lessons about dropout research and dropout prevention. Los Angeles, CA: The Civil Rights Project. Retrieved from <http://civilrightsproject.ucla.edu/research/k-12-education/school-dropouts/revisiting-the-myth-of-the-texas-miracle-in-education-lessons-about-dropout-research-and-dropout-prevention-walt-haney>
- Hayes, H. M. III. (1997). *An analysis of the interaction of prior knowledge, text presentation, and the purpose for reading text on recall*. (Doctoral dissertation). Retrieved from ProQuest Digital Dissertations. (UMI 9822607).
- Heier, S. (2011). *The relationship between standardized test scores of socioeconomically disadvantaged students in Title I and non-Title I schools*. (Doctoral dissertation). Retrieved from ProQuest Digital Dissertations. (UMI 3482830).
- Heubert, J. P., & Hauser, R. M. (Eds.). (1999). *High stakes: Testing for tracking, promotion, and graduation*. Washington, DC: National Academy Press.
- Ikegulu, T. N. (2004). *The impacts of demographic factors in predicting student performance on a state reading test*. Retrieved from ERIC database. (ED509320)
- Jary, D., & Jary, J. (2000). *Collins dictionary of sociology* (3rd ed.). Glasgow, Scotland: Harper Collins.
- Jaska, P., Hogan, P., & Wen, Z. (2009). Academic accountability in Texas public schools: 2003~2007. *Contemporary Issues in Education Research*, 2 (5), 59-64.
- King, M., Fagan, B., Bratt, T., & Baer, R. (1987). ESL and social studies integration. In J. Crandall (Ed.), *ESL through content area instruction: Mathematics, science, social studies* (pp. 89-119). Arlington, VA: Center for Applied Linguistics.

- Kobayashi, M. (2002). Method effects on reading comprehension test performance: Text organization and response format. *Language Testing*, 19 (2), 193-220.
- Koretz, D. (2002). Limitations in the use of achievement tests as measures of educators' productivity. *Journal of Human Resources*, 37 (4), 752-777.
- Laca, E. A. (2012). Chapter 15: Multivariate analysis of variance [Word Document]. Retrieved from Lecture Notes Online Web site:
<http://www.plantsciences.ucdavis.edu/agr206/Lectures.htm>
- Lea, S. (1997). *Multiple regression: basic concepts and procedures*. Retrieved from Lecture Notes Online Website:
<http://people.exeter.ac.uk/SEGLEa/psy2005/basicmlt.html>
- Lewis-Moreno, B. (n.d.). The ESL program in SAISD. Retrieved from
<http://www.saisd.net/admin/curric/bilingual/fageslsasisd.html>
- Lorch, R. F. Jr. & Chen, A. H. (1986). Effects of number signals on reading and recall. *Journal of Educational Psychology*, 78, 263-270.
- Manzo, K. K. (2005, March 16). Social studies losing out to reading, math. *Education Week*, 24 (27), 16-17.
- Menken, K. (2000). What are the critical issues in wide-scale assessment of English language learners? Washington, D.C: National Clearinghouse for Bilingual Education.
- Narveson, R. M. (2001). *Bulleted points and typographic cues: Effects on recall and recall order*. (Doctoral dissertation). Retrieved from ProQuest Digital Dissertations. (UMI 3008717).
- Nathans, L. L., Oswald, F. L., & Nimon, K. (2012). Interpreting multiple linear regression: A guidebook of variable importance. *Practical Assessment, Research & Evaluation*, 17 (9), 1-19. Retrieved from
<http://pareonline.net/getvn.asp?v=17&n=9>
- National Clearinghouse for English Language Acquisition (NCELA). (2011). *The growing numbers of English language learner students, 1998/99/2008/09*. Washington, D.C: NCELA. Retrieved from
www.ncela.gwu.edu/files/uploads/9/growingLEP_0809.pdf
- National Council of Teachers of Mathematics (NCTM). (2000). *Results from the seventh mathematics assessment of the National Assessment of Educational Progress*. Reston, VA: NCTM.

- Nichols, J. (2003). Prediction indicators for students failing the state of Indiana high school graduation exam. *Preventing School Failure, 47* (3), 112-120.
- Pallant, J. (2007). *SPSS survival manual* (3rd ed.). New York, NY: McGraw Hill.
- Phelps, R. (2003). *Kill the messenger: The war on standardized testing*. New Brunswick, NJ: Transaction Publishers.
- Pyron, M. (2007). "I hear you, but I don't understand you": *The effects of peer tutoring for helping secondary ESL students achieve academic success*. (Doctoral dissertation). Retrieved from ProQuest Digital Dissertations. (UMI 304836128).
- Quezada, R. J. (2008). *The impact of principals' teacher certification on middle school student achievement in an urban South Texas school district*. (Doctoral dissertation). Retrieved from ProQuest Digital Dissertations. (UMI 3401420).
- Quinn, G. P., & Keough, M. J. (2002). *Experimental design and data analysis for biologists*. New York, NY: Cambridge University Press.
- Reilly, T. (1988). *ESL through content area instruction*. Washington, D.C: ERIC Clearinghouse on Languages and Linguistics.
- Reiss, P. P. (2005). *Causal models of item format and gender-related differences in performance on a large-scale mathematics assessment for grade three to grade ten*. (Unpublished doctoral dissertation). University of Hawaii at Manoa, Hawaii.
- Roane, W. (2008, July). Perspectives from Latin America: International lessons for U.S. English language learners. Paper presented at the 11th International Congress on Mathematical Education, Monterrey, Mexico. Retrieved from <http://www.eric.ed.gov/PDFS/ED502466.pdf>
- Segall, A., & Gaudelli, W. (2007). Reflecting socially on social issues in a social studies methods course. *Teaching Education, 18* (1), 77-92.
- Seixas, P. (1993). Historical understanding among adolescents in a multicultural setting. *Curriculum Inquiry, 23* (3), 301-327.
- Shin, S. (2002). Effects of subskills and text types on Korean EFL reading scores. *Second Language Studies, 20* (2), 107-130.
- Short, D. J. (1994). Expanding middle school horizons: Integrating language, culture, and social studies. *TESOL Quarterly, 28* (3), 581-608.

- Silva, P. (2010). *School and student characteristics of non-graduating seniors due to failure of TAKS tests in the Western Independent School District*. (Doctoral dissertation). Retrieved from ProQuest Digital Dissertations. (UMI 3433514).
- Sirin, S. (2005). Socioeconomic status and academic achievement: A meta-analytic review of research. *Review of Educational Research*, 75, 417-453.
- Slaughter, S. D. (2007). *Demographic profiles associated with academic performance for third grade students in North Forest and Aldine Independent School Districts in Texas*. (Doctoral dissertation). Retrieved from ProQuest Digital Dissertations. (UMI 3270397).
- Smith, E. (2008). Pitfalls and promises: the use of secondary data analysis in educational research. *British Journal of Educational Studies*, 56 (3), 323-339.
- Smith, N. (2012). *Directing curriculum through standards: A content analysis of the 2010 Texas state social studies standards*. (Doctoral dissertation). Retrieved from ProQuest Digital Dissertations. (UMI 3507749).
- StatSoft, Inc. (2012). How to find relationship between variables, multiple regression. In *Electronic Statistics Textbook*. Retrieved from <http://www.statsoft.com/textbook/multiple-regression/#cleast>
- Stewart, L. (2009). Achievement differences between large and small schools in Texas. *The Rural Educator*, 30 (2), 20-28.
- Strayhorn, T. L. (2009). Accessing and using national databases. In T. J. Kowalski & T. J. Lasley II (Eds.), *Handbook on data-based decision making in education* (pp. 105-122). New York, NY: Routledge.
- Tabachnick, B. G., & Fidell, L. S. (2012). *Using multivariate statistics* (6th ed.). Boston, MA: Allyn and Bacon.
- Texas Code on Education. (2009). Chapter 4: Public education mission, objectives, and goals. Retrieved from <http://law.justia.com/codes/texas/2009/education-code/title-2-public-education/chapter-4-public-education-mission-objectives-and-goals/>
- Texas Education Agency. (2003a). *Academic Excellence Indicator System 2002-03* [PDF document]. Retrieved from <http://www.tea.state.tx.us/perfreport/aeis/2003/state.html>.
- Texas Education Agency. (2003b). Technical Digest 2002-2003. Retrieved from <http://www.tea.state.tx.us/student.assessment/techdigest/yr0203/>

- Texas Education Agency. (2004a). Texas Assessment of Knowledge and Skills information booklet: Social studies grade 8. Retrieved from <http://www.tea.state.tx.us/student.assessment/taks/infobooks/>
- Texas Education Agency. (2004b). Texas Assessment of Knowledge and Skills information booklet: Social studies exit level. Retrieved from <http://www.tea.state.tx.us/student.assessment/taks/infobooks/>
- Texas Education Agency. (2006a). *Academic Excellence Indicator System 2005-06* [PDF document]. Retrieved from <http://www.tea.state.tx.us/perfreport/aeis/2006/state.html>.
- Texas Education Agency. (2006b). *Data file format with student item analysis* [Data file]. Retrieved from www.tea.state.tx.us/WorkArea/linkit.aspx?LinkIdentifier=id&ItemID=2147497322&libID=2147497319
- Texas Education Agency. (2006c). Technical Digest 2005-2006. Retrieved from <http://www.tea.state.tx.us/student.assessment/techdigest/yr0506/>
- Texas Education Agency. (2007a). Texas Assessment of Knowledge and Skills study guide: Social studies grade 8. Retrieved from http://www.tea.state.tx.us/index3.aspx?id=60&menu_id=793
- Texas Education Agency. (2007b). Texas Assessment of Knowledge and Skills study guide: Social studies grade 11. Retrieved from http://www.tea.state.tx.us/index3.aspx?id=60&menu_id=793
- Texas Education Agency. (2008). Technical Digest 2007-2008. Retrieved from <http://www.tea.state.tx.us/student.assessment/techdigest/yr0708/>
- Texas Education Agency. (2009a). *Academic Excellence Indicator System 2008-09* [PDF document]. Retrieved from <http://www.tea.state.tx.us/perfreport/aeis/2009/state.html>.
- Texas Education Agency. (2009b). Glossary of terms, 2007-2008 division of accountability research. Retrieved from <http://ritter.tea.state.tx.us/acctres/gloss0708.html>
- Texas Education Agency. (2009c). Majority of students meet TAKS promotion requirements. Retrieved from <http://www.tea.state.tx.us/index4.aspx?id=4504>

- Texas Education Agency. (2009d). Technical Digest 2008-2009. Retrieved from <http://www.tea.state.tx.us/student.assessment/techdigest/yr0809/>
- Texas Education Agency. (2009e). *Title III language program definitions* [PDF file]. Retrieved from http://www.elltx.org/docs/LanguageProgramDef_Mar2012.pdf
- Texas Education Agency. (2010). Technical Digest 2009-2010. Retrieved from TEA Website: <http://www.tea.state.tx.us/student.assessment/techdigest/yr0910/>
- Texas Education Agency. (2011a). Enrollment in Texas public schools 2010-2011. Retrieved from <http://www.tea.state.tx.us>
- Texas Education Agency. (2011b). Social studies TEKS. Retrieved from <http://www.tea.state.tx.us/index2.aspx?id=3643>
- Texas Education Agency. (2011c). TAKS performance standards. Retrieved from <http://www.tea.state.tx.us/student.assessment/taks/pstandards/>
- Texas Education Agency. (2012a). Limited English proficiency initiatives. Retrieved from http://www.tea.state.tx.us/index4.aspx?id=5081&menu_id=814
- Texas Education Agency. (2012b). *State of Texas Assessments of Academic Readiness (STAARTM): Questions and answers (Q&As)* [PDF file]. Retrieved from <http://www.tea.state.tx.us/student.assessment/staar/>
- Texas Education Agency. (2012c). Texas English language learner portal. Retrieved from <http://www.elltx.org/index.html>
- Texas Education Code. (n.d.). Chapter 29.052: Definitions. Retrieved from <http://www.statutes.legis.state.tx.us/Docs/ED/htm/ED.29.htm#B>
- U. S. Department of Education. (2002). *The No Child Left Behind Act of 2001* [PDF file]. Retrieved from <http://www2.ed.gov/nclb/overview/intro/execsumm.html>
- U.S. Department of Education. (n.d.). Improving basic programs operated by local educational agencies (Title I, Part A). Retrieved from <http://www2.ed.gov/programs/titleiparta/index.html>
- van Langen, A., Bosker, R., & Dekkers, H. (2006). Exploring cross-national differences in gender gaps in education. *Educational Research and Evaluation*, 12 (2), 155-177.
- Westheimer, J., & Kahne, J. (2003). Democracy and civic engagement, reconnecting education to democracy: Democratic dialogues. *Phi Delta Kappan*, 85 (1), 8.

White, K. R. (1982). The relation between socioeconomic status and academic achievement. *Psychological Bulletin*, 91 (3), 461-481

Zamarripa, L. (2009). *Factors affecting student achievement in mathematics in select Texas high schools*. (Doctoral dissertation). Retrieved from ProQuest Digital Dissertations. (UMI 3400331).