

Copyright

by

Lenora Laverpool-Leatherwood

December 2013

DOES BLENDED INSTRUCTION MATTER FOR ENGLISH-LANGUAGE
LEARNERS? A STUDY ASSESSING THE INFLUENCE OF INSTRUCTIONAL
METHOD ON LANGUAGE ACQUISITION

A Dissertation Proposal Presented to the
Faculty of the College of Education
University of Houston

In Partial Fulfillment
of the Requirements for the Degree

Doctor of Education

by

Lenora Laverpool-Leatherwood

December 2013

DOES BLENDED INSTRUCTION MATTER FOR ENGLISH-LANGUAGE
LEARNERS? A STUDY ASSESSING THE INFLUENCE OF INSTRUCTIONAL
METHOD ON LANGUAGE ACQUISITION

A Dissertation for the Degree
Doctor of Education

by

Lenora Laverpool-Leatherwood

Approved by Dissertation Committee

Dr. Laveria Hutchinson, Chairperson

Dr. Lee Mountain, Committee Member

Dr. Margaret Watson, Committee Member

Dr. Miguel Ramos, Committee Member

Dr. Robert McPherson, Dean
College of Education

December 2013

Acknowledgements

The day I stepped into my first education course at Houston Baptist University, I knew with certainty that one day I would teach future teachers. I thank God for the vision, that it never perished, and that He sustained me on the journey.

To my advisor, Dr. Laveria Hutchinson, thank you for your guidance. Thank you to my committee members, Dr. Lee Mountain, Dr. Miguel Ramos, and Dr. Margaret Watson, who challenged, assisted, and encouraged me along the way. I invited each of you to be a part of the committee because throughout the initiative, each of you inspired me and helped equip me for the journey. Please know that you are appreciated.

Rebecca “Becky” Perez, thank you for being great and what you do and then some. It is because of you that I met my graduation goal. You made the difference. To my family and friends who cheered me on, your encouragement and prayers carried me along the way. A special thank you to Dr. Katharine Garcia. Thank you for your listening ear, brilliant ideas, and willingness to always help. My friend, you made the journey easier. Also, to Dr. Walter Jackson, thanks for being my first editor. Brenda Caceres, my former ESL tutee, you went from struggling to pass the exit level ELA TAKS test to writing a novel. You inspire me.

To my mother, Shirley Wright Ekhaton, thank you for being my first role model and the person who taught me to love reading and to love learning. Maybe this means “I am a smart cookie.” If so it is only because you provided the right ingredients. Dearest brother Aaron, thank you for always checking in and for being the project manager for the graduation festivities. More often than not I was doing school work when you called, but I appreciate that you kept calling. Byron and Brenton, you kept me grounded in the

real world throughout the process. Whether it was surprise visits, a few rounds of tennis, or snack breaks, thank you my brothers for the much needed breaks.

Last but not least, a special thank you to my husband, Anthony Bryant Leatherwood. Thank you for your encouragement and support throughout this journey. I am appreciative that you patiently put our other goals on hold so that I could embark on this journey. I love you, appreciate you, and thank God for you.

DOES BLENDED INSTRUCTION MATTER FOR ENGLISH-LANGUAGE
LEARNERS? A STUDY ASSESSING THE INFLUENCE OF INSTRUCTIONAL
METHOD ON LANGUAGE ACQUISITION

An Abstract
of a Dissertation Presented to the
Faculty of the College of Education
University of Houston

In Partial Fulfillment
of the Requirements for the Degree

Doctor of Education

by

Lenora Laverpool-Leatherwood

December 2013

Leatherwood, Lenora. "Does Blended Instruction Matter for English-Language Learners? A Study Assessing the Influence of Instructional Method on Language Acquisition." Unpublished Doctor of Education Dissertation, University of Houston, December 2013.

Abstract

In the 21st century, Texas schools continue to experience exponential growth in their numbers of English language learners (ELLs). At the secondary level, high school ELLs are expected to take and pass the State of Texas Assessment of Academic Readiness End-of-Course (STAAR/EOC) exams as part of their cumulative graduation requirements. Additionally, ELLs' performance on the STAAR exams at all grade levels significantly impacts their schools' and districts' adequate yearly progress (AYP). Subsequently, many schools and districts in the state continue to explore innovative ways to help ELLs effectively and efficiently acquire the academic English they need to succeed in school. This study examined a high school that piloted English as a Second Language (ESL) English and Reading courses delivered through a blended method of instruction in order to determine if the method increases students' language acquisition at a greater rate than students taking ESL English and Reading courses delivered via a traditional method of instruction as measured by Texas English Language Proficiency Assessment System (TELPAS) reading tests.

Results indicate that there were no statistically significant differences in TELPAS 2013 reading scores between the group that received the blended method of instruction and the group that received the traditional method of instruction. The same was true in terms of STAAR English II reading scores, the differences between the groups were not statistically significant. ANOVA and ANCOVA did indicate a significant effect of TELPAS 2012 scores on TELPAS 2013 scores. Likewise there was a significant effect

of STAAR English I reading scores on STAAR English II reading scores. Further, results indicated that there were significant differences in TELPAS 2013 reading scores based on instructional level, age, STAAR English I reading scores, and STAAR English II reading scores, but there were not significant differences based on gender and years in U.S. schools. As for STAAR English II reading scores, there were significant differences in scores based on instructional level, years in U.S. schools, and TELPAS 2012, yet no significant differences based on gender and age

Table of Contents

Chapter	Page
Chapter I Introduction.....	1
Need for the Study	8
Statement of the Problem.....	9
Purpose of the Study	10
Research Questions.....	10
Summary and Significance of the Study.....	10
Definition of Terms.....	11
Chapter II Review of the Literature	14
Introduction.....	14
Second Language Acquisition	15
Best Practices in ESL Education.....	25
Blended Instruction.....	37
Chapter III Methodology	45
Introduction.....	45
Research Questions.....	45
Hypothesis.....	46
Sample Selection.....	46
The district	46
The high schools.	49

The teachers.	51
Traditional instruction model.....	55
Blended instruction model	56
The students.	58
Instrumentation	63
Data Analysis Procedures	67
Dependent Variables.....	67
Demographic Variables	68
Chapter IV Results.....	69
Descriptive Statistics Results.....	70
Gender.....	70
Home Language.....	70
Instructional Level.	70
Age.....	71
Refugee Status.	72
Years in U.S. Schools	72
STAAR English I Reading (Grade 9).	73
STAAR English II Reading (Grade 10).....	73
TELPAS 2012.....	74
TELPAS 2013.....	75

Analysis of Variance and Analysis of Covariance Results.....	75
TELPAS 2013	75
STAAR English II Reading (Grade 10).....	82
Summary	90
Chapter V Conclusions, Interpretations, Recommendations	92
Introduction.....	92
Language Acquisition as Measured by TELPAS 2013 Interpretations	93
TELPAS 2013 and the Independent Variables.	96
Language Acquisition as Measured by STAAR English II Interpretations.....	101
STAAR English II Reading and the Independent Variables	103
Limitations	106
Implications for Practice	107
Recommendations for Future Research	108
Summary	109
References.....	110
Appendix A TELPAS Ratings	120
Appendix B Traditional Instructional Method Lesson Plans.....	122
Appendix C Blended Instructional Method Lesson Plans	129
Appendix D Descriptive Statistics for All Variables.....	131
Appendix E Home Language Frequencies	135

Appendix F Tests of Between Subjects Dependent Variable TELPAS 2013	138
Appendix G ANOVA Descriptive Statistics by Group TELPAS 2013.....	142
Appendix H Tests of Between Subjects STAAR10	146
Appendix I ANOVA Descriptive Statistics by Group STAAR 10.....	150
Appendix J STAAR 10 Post Hoc Test Years in U.S. Schools	154

List of Tables

Table	Page
1. Blended Instruction Continuum	40
2. Models of Blended Instruction	41
3. High School Demographics 2012-2013	50
4. High School 1 Teacher A Treatment Group	61
5. Comparison Group	63
6. Tests of Between-Subjects Effects TELPAS 2013 Type of Group	78
7. Games-Howell Post Hoc Test TELPAS 2013/ Instructional Level	79
8. ANCOVA TELPAS 2013 Covariate TELPAS 2012	82
9. Test of Between-Subjects STAAR 10 Type of Group	84
10. Tukey HSD Post Hoc Test STAAR 10/ Instructional Levels	86
11. Tukey HSD Post Hoc Test STAAR 10/TELPAS 2013	89
12. ANCOVA STAAR 10 Covariate STAAR 9	90

List of Figures

Figure		Page
1	The Dimensions of Language Proficiency	21

Chapter I

Introduction

The No Child Left Behind (NCLB) Act of 2002 ushered in a climate of change in America's public education system. Most notably, this groundbreaking education legislation brought academic accountability to the forefront of education for the first time, thus making states, districts, and schools accountable for the adequate yearly progress (AYP) of students based upon their participation in and performance on high-stakes standardized tests. Performance data of students reported for AYP includes general education students as well as specific populations like special education students and English-language learners (ELLs) who had not traditionally received academic primacy. Instead, districts and schools tended to spend their academic resources, such as interventions and instructional materials, on general education students because they were the only students who participated in standardized tests. As a result of the accountability measures and the respective penalties of NCLB, federal, state, and local attention has been given to ELLs, a rapidly growing student group in U.S. schools.

The National Clearinghouse for English Language Acquisition reports that the number of ELLs enrolled in U.S. schools rose from 3.5 million during the 1998-1999 school year to 5.3 million in the 2008-2009 school year, an increase of over 50 percent (National Clearinghouse for English Language Acquisition, 2010). As of 2011, one in every nine students was identified as an ELL, and several states including South Carolina, Kentucky, and Virginia reported significant increases in the numbers of ELLs ranging from 100 to 200 percent respectively (National Clearinghouse, 2010). Although many states are working with record numbers of ELLs for the first time, states like

California and Texas, which have worked with immigrant students for years, continue to serve the greatest number of ELLs (National Clearinghouse, 2010). Specifically, Texas' ELL enrollment increased from 570,000 in 2001 to over 800,000 in 2010, making it the state with the second highest ELL enrollment behind California which services over 1.1 million ELLs in its PK-12 public schools (National Clearinghouse, 2010). As the number of ELLs steadily rises, the challenge of meeting their academic needs has become critical when bearing in mind the onerous demands of AYP. Further the task becomes conjectural as Texas transitions to a new system of accountability that is consistently evolving because all aspects of it have not been finalized.

Texas public schools are in a state of adjustment as the Texas Education Agency (TEA) transitions to a new accountability system that represents a merger of the once separated state and federal accountability systems. Prior to the new accountability system proposal, all public school districts and campuses were evaluated annually and labeled as either meeting or missing AYP (Texas Education Agency, 2010). The overall rating was based on graduation rates, attendance rates, students' participation on standardized tests in math and English/Language Arts, and students' academic performance on the math and English/Language Arts standardized tests (Texas Education Agency, 2010). Comparably, the design of the new accountability system is a performance index framework with indicators grouped into four indices: student achievement, student progress, closing performance gaps, and postsecondary readiness (Texas Education Agency, 2013b). Under the new performance index accountability system, districts and schools will be assigned one of the following ratings: met standard,

met alternative standard, or improvement required (Texas Education Agency, 2013b). In addition, the plan for spring 2014 and beyond is for the state to assign accountability rating labels of A, B, C, D, and F (Texas Education Agency, 2013b).

Similar to the outgoing AYP structure, the new accountability system requires the disaggregation of data by various student groups assuming the group meets the minimum size requirement as specified by each index with the exception of index one (Texas Education Agency, 2013a). Index one includes all students in all subgroups regardless of size. The minimum requirement for a subgroup to be included in indices two and three is if 25 tests are administered to any given subgroup across all subjects (Texas Education Agency, 2013a). For example, a school may have five students in a subgroup who each take five exams (such as reading, writing, math, science, and social studies) and therefore their data is included in the indices. Index four has a minimum requirement of 25 students in the subgroup for it to be used in graduation rate calculations (Texas Education Agency, 2013a). The performance indices of the new accountability system include performance of *all* students and the following ten student groups: economically disadvantaged, English-language learners, special education, African American, American Indian, Asian, Hispanic, Pacific Islander, white, and Two or More Races (Texas Education Agency, 2013a). It is important to note that students can be included in more than one sub-group; for example, a student may be Two or More Races, economically disadvantaged, an English-language learner, and in the “all” subgroups. Subsequently, a student’s performance on the state exams may impact a school’s AYP in

multiple subgroups and multiple indices. Therefore ELLs' performance on the state tests is critical as schools and districts strive to meet AYP standards.

As of the 2013-2014 school year, the more rigorous State of Texas Assessment of Academic Readiness (STAAR) tests replaced the Texas Assessment of Knowledge and Skills (TAKS) tests at all grade levels, adding another dimension of complexity for districts and schools striving to help ELLs pass the state's exams. The STAAR was first implemented during the 2011-2012 school year for third, fourth, seventh, eighth, and ninth graders, and was implemented for tenth graders during the 2012-2013 school year. Furthermore, as part of their graduation requirements, ELLs must take and pass the STAAR or the linguistically accommodated test, STAAR L (Texas Education Agency, 2013a). ELLs who have not attained an advanced high rating on the Texas English Language Proficiency Assessment System (TELPAS) in reading and who have been enrolled in U.S. schools for three years or less, or five years or less for unschooled refugees and asylees, may take the STAAR L in math, science, and/or social studies as part of their graduation requirement if they require "moderate to substantial linguistic accommodations to understand the English used on STAAR mathematics, science, and social studies assessments" (Texas Education Agency, 2013a). However, students may take STAAR L while in one grade yet may no longer qualify for the STAAR L the following school year (Texas Education Agency, 2013a). Additionally, if a student takes the STAAR L in the spring of his/her ninth grade year and fails the exam and has to retake it, he/she may no longer qualify to take the STAAR L and therefore has to take the regular STAAR exam the following spring (Texas Education Agency, 2013a). Exceptions

to this relate to the STAAR reading and STAAR writing assessments because there are no linguistically accommodated versions of these exams. Instead, ELLs take the same exams as their native-English-speaking peers. If, however students qualify to take the STAAR L in math, science, and/or social studies, then they also qualify for a special provision for the ninth grade English STAAR exams in reading and writing (Texas Education Agency, 2013a). When students receive a special provision, it means that they do not have to retake the exams as part of their graduation requirements if they fail them. There are however no special provisions for the tenth grade English STAAR exams in reading and writing; meaning ELLs must take and pass the exams in order to graduate (Texas Education Agency, 2012b). Whether ELLs take the STAAR or the STAAR L, their scores are used for state and federal accountability purposes.

Adequate Yearly Progress data shows that ELLs lag far behind their native-English speaking peers when it comes to performance on standardized tests (Flores, Batalova, & Fix, 2012; Zacher-Pandya, 2011). Unfortunately, ELLs are placed at a distinct disadvantage when their scores are compared to those of native-English speakers because ELLs who are still working to acquire academic English typically do not have the same academic language skills in English as native-English speakers (Ariza, 2009). The reason for this is because acquiring a first language (L1) has been described as both a complex and a lifelong process (Collier, 1989). Students identified as native speakers continue to acquire the English language and its “subtleties in pragmatics as well as the constantly changing patterns in language that affect everyday oral and written communication with others” (Collier, 1989, p.12) well into their adult lives. Since native

speakers are constantly adding to their own language proficiency, getting ELLs to be at the same English proficiency as their native-English speaking peers is what Thomas refers to as chasing a moving target (as cited in Collier, 1989, p. 4). All the more, acquiring a second language (L2), i.e. English, is described as being equally challenging (Collier, 1989). When first acquiring L2, newcomers learn what Cummins (1981) labeled as Basic Interpersonal Communication Skills (BICS) which is functional language. According to research, it takes one to three years for immigrants to acquire social language (Collier, 1995; Cummins 1981; Hakuta, Butler, & Witt, 2000; Thomas & Collier, 1997). Acquiring what Cummins (1981) labeled as Cognitive Academic Language Proficiency (CALP), the academic language of learning in content areas, takes longer to acquire than does BICS (Collier, 1995; Cummins 1981; Thomas & Collier, 1997). Studies have found that it takes ELLs who have at least two to three years of education in L1 anywhere from five to seven years minimum to acquire academic proficiency in L2 (Collier, 1995; Collier & Thomas, 1989; Cummins, 1981; Hakuta et al., 2000). Further, ELLs with no formal education in their L1 need at least seven to ten years to acquire academic English (Collier, 1995; Cummins, 1981; Thomas & Collier, 1997). English-language learners' success on high-stakes tests hinges on their acquisition of academic English. However, since districts and schools are not afforded the luxury of waiting at least five years to administer standardized tests to incoming ELLs, educators are faced with the critical challenge of helping immigrant students acquire academic English both effectively and efficiently if they are to meet AYP standards.

One approach that may prove effective in helping ELLs acquire English is the use of a blended instructional method. The most general definition of blended instruction is a combination of face-to-face and online instruction (Horn, Staker, & Innosight, 2011; Picciano & Seaman, 2009; Watson, 2008; & Watson, Murin, Vashaw, Gemin, & Rapp, 2010). Watson (2008) posits that “the blended approach combines the best elements of online and face-to-face learning” and “is likely to emerge as the predominant model of the future” (p. 3). Although online learning has traditionally been more prevalent in higher education, it is becoming more pervasive in K-12 schools. In 2000, there were over 45,000 K-12 students enrolled in some form of online learning, and by 2009, that number had increased to over three million (Horn et al., 2011). Researchers believe that a part of the reason for the rise in K-12 online enrollment is increased adoption of the blended instructional model (Horn et al., 2011).

For the purpose of this study, blended instruction is defined as instruction that is comprised significantly of an online component and is held in a computer lab where students meet every day with a teacher who provides a portion of the instruction (Watson, 2008). Perhaps the most advantageous benefit of blended learning in terms of ELLs is that it allows educators to individualize instruction based on students’ instructional needs (Horn et al., 2011). Both learning and instruction are personalized for students and allows them to work at their own pace (Horn et al., 2011). Further advantages of blended instruction include that it offers students an enriched experience using technology as well as opportunities to interact with their teacher and their peers (Watson et al., 2010). As a result of its proclivity to help teachers individualize instruction, its ability to grant

students 21st century skill development, and its capacity to provide students opportunities to interact with their teacher and their peers, blended instruction may be an educational approach that positively impacts ELLs' language acquisition.

Preparing ELLs to pass high-stakes tests is a complex endeavor that is predicated by their acquisition of academic language. This task falls to educators who are faced with the challenge of helping ELLs acquire academic English. Therefore, districts and teachers are searching for optimal instructional approaches to help ELLs acquire academic English both effectively and efficiently.

Need for the Study

The substantive role of academic language acquisition in ELLs' success on academic tasks such as standardized tests has been supported by numerous researchers (Collier, 1995; Cummins, 1981; Thomas & Collier, 1989; Zacher-Pandya, 2011). Therefore, as districts and schools have the responsibility to help ELLs pass high-stakes exams, they must adopt sound instructional practices that assist students in learning English. One district implemented a blended instructional model with ELLs to achieve this goal.

Although the blended instructional approach has its origins in adult education, it is rapidly becoming more widespread in K-12 schools (Zacher-Pandya, 2011). As the popularity of blended instruction grows, it is necessary to understand its impact on ELLs since they account for one in every nine students in American schools (National Clearinghouse, 2010). More importantly, it is imperative to understand the impact of

blended instruction on ELLs' language acquisition since their success on standardized tests hinges upon it.

This study is needed because it proposes to examine how blended instruction, a growing method used in K-12 schools, impacts the language acquisition of ELLs.

Although there is research available that compares the outcomes of online learning versus traditional learning, the research is limited in terms of comparing traditional instruction and blended instruction in K-12 education when it comes to ELLs. Furthermore, none of the studies address the impact on language acquisition. This study examines the blended instructional approach and its impact on high school ELLs' language acquisition.

Statement of the Problem

As the number of ELLs in U.S. schools continues to increase, so does the importance and the impact of meeting their academic needs. While the challenge is new to several states like South Carolina and Virginia, others like Texas have been addressing the educational needs of ELLs for well over a decade. It has been the accountability measures of NCLB and their respective penalties that have brought educating special populations like ELLs to the academic forefront. States, districts, and schools striving to meet AYP standards are searching for instructional practices that will help ELLs pass standardized tests. Since ELLs passing standardized tests is contingent upon their acquisition of academic English, districts and teachers are searching for instructional practices that will help students acquire academic English both effectively and efficiently.

Purpose of the Study

The purpose of the study is to compare the effect on language acquisition of ELLs taking a blended ESL English or ESL reading class to the language acquisition of ELLs taking a traditional ESL English or ESL reading class. Language acquisition will be measured using the TELPAS reading and STAAR reading assessments. The study explores the relationship between blended ESL instruction and language acquisition. It helps answer the question about the impact of blended instruction on students' academic language acquisition.

Research Questions

This study investigates the relationship between blended instruction and the language acquisition of ELLs. Specifically, this study explores the following questions:

1. Is there a difference in the language acquisition of ELLs who take a blended ESL English class or a blended ESL reading class in comparison to ELLs who take a traditional ESL English or ESL reading class as measured by the Texas English Language Proficiency Assessment System reading exam?
2. Is there a difference in the language acquisition of ELLs who take a blended ESL English class or a blended ESL reading class in comparison to ELLs who take a traditional ESL English or ESL reading class as measured by the State of Texas Assessment of Academic Readiness English II reading exam?

Summary and Significance of the Study

The growing numbers of ELLs in Texas public schools along with the legislation of NCLB have combined to create an educational crisis. Since states, districts, and

schools are held accountable for the performance of ELLs on high-stakes tests, they are working perspicaciously to identify the most efficient and effective ways to help ELLs acquire the academic English necessary to perform on these tests. This study is significant because it adds to the knowledge base about instructional modalities for ELLs. Furthermore, it looks at a blended instructional method and places it in the broader context of helping ELLs acquire academic English.

Definition of Terms

Blended Instructional Method: An instructional approach that is comprised mostly/significantly of an online component and is held in a computer lab where students meet every day with a teacher who provides a portion of the instruction (Watson, 2008).

English Language Learners (ELLs): A student whose primary language is a language other than English and whose English language skills are such that the student has difficulty performing ordinary classwork in English (Texas Education Agency, 2013c).

English as a Second Language (ESL): Courses taught specifically for ELLs in English but with second language support (Texas Education Agency, 2013c).

L1: A students' native language; first language (Echevarria et al, 2013).

L2: A students' second language; target language (Echevarria et al., 2013).

Language Acquisition: The process of attaining the social, cultural, and cognitive designs of a language. Learning a new language (Cummins, 1981).

Language Proficiency Assessment Committee (LPAC): The decision making committee for English language learners that makes decisions on an individual basis for

each ELL in accordance with the procedures established by TEA (Texas Education Agency, 2012c).

Limited English Proficient (LEP): Students of limited English proficiency whose primary language is a language other than English and whose English language skills are such that students have difficulty performing ordinary classwork in English (Texas Education Agency, 2013c).

Sheltered English: English course in which “Teachers deliver grade-level objectives for the course to English learners through modified instruction that makes the information comprehensible to the students while promoting their academic English development” (Echevarria et al., 2008, p. 15).

Sheltered Instruction Observation Protocol Model (SIOP): “A research-based and validated model of sheltered instruction that helps teachers plan and deliver lessons that allow English learners to acquire academic knowledge as they develop English-language proficiency” (Center for Applied Linguistics, 2013).

State of Texas Assessment of Academic Readiness (STAAR) and End of Course (EOC): Standardized tests developed by the state of Texas that assess students in reading, writing, math, science, and social studies. STAAR/EOC is used at the high school level to mean the STAAR assessments administered to high school students at the end of the course (Texas Education Agency, 2010).

Texas English Language Proficiency Assessment System (TELPAS): Test designed to assess the progress that limited English proficient (LEP) students make in learning the English language (Texas Education Agency, 2011b).

Traditional Instructional Method: Instructional method that includes the workshop model described below, the Madeline Hunter lesson design format, or a combination of the workshop model and Madeline Hunter lesson format.

Workshop Model: A method of teaching English that involves an opening structure, a mini-lesson, work time, and a debrief. In this framework students in a classroom are working on reading, writing, and talking (in some combination) for approximately two-thirds of the class period. The premise of the model is that whoever does the reading, writing, and talking is doing the thinking and learning is a consequence of thinking. During the worktime, teachers confer with students and determine what they know, can do, and still need in order to be successful (Bennett, 2007).

ChapterII

Review of the Literature

Introduction

The focus of this research is to compare the effect on second language acquisition of taking a blended ESL English or reading class versus taking a traditional ESL English or reading class. For the purpose of this study, the review of the literature was obtained through a structured criterion.

The studies collected for the literature review are empirical research, including both quantitative and qualitative areas. Information and data collected for the review of literature reside within a twenty year span; however, the literature review may include other prominent studies carried out prior to these timeline parameters. Also, the literature reflects samples from peer-reviewed journals and selected technical reports, and includes books and book chapters.

The journal searches were conducted using four large databases: Academic Search Premier, Education Search Complete, Educational Resources Information Center (ERIC), and Psych INFO. Then, the articles were sought utilizing the following key terms: second language acquisition, English-language learners, blended instruction, best practices and English-language instruction, and blended instruction and English-language learners.

Manual and computer searches were carried out in specific journals relevant to ELLs and best practices (for example, Teachers of English of Speakers of Other Languages) in order to find potential studies to include in the literature review. While some of the research entails an in-depth, area-specific analysis, the review of the

literature is best categorized as inclusive. Thus, the literature review highlights, denotes and includes pertinent research conducted with a broad subject focus, yet is guided by the stated criterion parameters already established within the frame of this study.

The material assembled centers around the major aspects of the proposed study: second language acquisition, best practices in ESL education, and blended instruction.

Second Language Acquisition

Over the years, there have been several theories of second language acquisition (SLA). Some of the earliest theories will be briefly described and the most significant to the proposed study will be more thoroughly explained.

One of the earliest theories of SLA is contrastive analysis which is based on behaviorism. It was proposed by Lado in 1957 and is the systematic study of two languages to identify the structural differences and similarities between L1 and L2 as the relationship between the two is purported to play a crucial role in learners' L2 production (Richard-Amato, 2003; Rustipa, 2011). The theory suggests that similarities between L1 and L2 promote language acquisition, while differences attribute to interference of L1 in the mastery of L2 (Richard-Amato, 2003; Rustipa, 2011). It posits that problems/errors can be predicted based on the differences between the languages. Theorists assert that L1 habits must be broken to attain L2 firmly. Mitchell and Myles (2004), among other linguists, disagree with the constrastive analysis theory. They argue that several research studies support that errors made in L2 acquisition are not due to interference from L1 (Mitchell & Myles, 2004). Therefore, they claim, contrastive analysis is not a valid predictor of L2 errors as errors L2 learners make while acquiring L2 cannot be traced to

differences between L1 and L2 (Mitchell & Myles, 2004). Further, linguists reject the theory because not all predicted errors in L2 production occur, and not all errors made are predicted.

Another SLA theory is error analysis hypothesis which is influenced by developmentalism, the belief that “learning develops in variable stages as learners interact with the environment” (Richard-Amato, 2003, p. 36). It was proposed by Corder in 1967 as an alternative to contrastive analysis theory (Richard-Amato, 2003; Rupita, 2011). The theory involves looking at errors made by the L2 learners and trying to identify why the errors are made, what they suggest, and if they mean that the learner is doing something incorrect or if the learner has acquired a rule and is indeed progressing in L2 acquisition (Richard-Amato, 2003). Contrary to contrastive analysis, L1 is seen as beneficial to the process of developing L2. Errors are considered necessary in L2 acquisition and allegedly account for all L2 errors. Opponents of the theory argue that there is no empirical data to substantiate that all errors made by learners as they acquire L2 are a result of interference of L1 (Richard-Amato, 2003). Further, they note that the theory does not provide information about the language learners that impacts L2 production such as educational background and age (Richard-Amato, 2003).

A third theory of SLA is interlanguage development which is the progression of a learner’s language as it moves toward L2 production (Richard-Amato, 2003). It is theorized that as students develop L2, they rely less and less on L1 (Richard-Amato, 2003). Subsequently, as students move toward L2 proficiency they develop an interim language referred to as interlanguage development which is neither L1 nor L2 (Richard-

Amato, 2003). Instead, the L2 learner constructs abstract linguistic rules while progressing toward fluency in L2 (Richard-Amato, 2003). Moreover, the process for L2 acquisition is considered similar to that of L1 acquisition in that there is a systematic development of syntax, pragmatics, and semantics (Richard-Amato, 2003).

Notably, another theory of SLA was proposed by Stephen Krashen. Krashen's (1985) theory consists of five hypotheses: acquisition-learning, natural order, monitor, input, and affective filter. The acquisition-learning hypothesis differentiates between acquisition and learning; acquisition is described as a subconscious process while learning is described as a conscious process. Natural order avows that the rules of language are acquired in a predictable order in which some rules are acquired early in L2 development and others are acquired later. The monitor hypothesis explains how acquisition and learning are used in L2 production. Krashen (1985) denotes that the ability to produce L2 comes from the subconscious acquired competence, and learning serves as a monitor used to make corrections to what is acquired. The input hypothesis asserts that L2 is acquired by receiving comprehensible input. He posits that the only way for L2 to develop is when learners understand input. Affective filter is the fifth and final hypothesis. It is described as a mental block that may interfere with L2 learners fully using the comprehensible input they receive to acquire L2. The affective filter may be impacted negatively by issues such as a fear of failure, lack of motivation, and low self-confidence. Conversely, it may be positively impacted by attributes such as high motivation and high self-esteem. In order to acquire L2, learners must have their

affective filters lowered. Krashen (1985) best summarizes the interaction among the hypotheses by noting:

People acquire second languages only if they obtain comprehensible input and if their affective filters are low enough to allow the input 'in'. When the filter is 'down' and appropriate comprehensible input is presented (and comprehended), acquisition is inevitable. It is, in fact, unavoidable and cannot be prevented- the language 'mental organ' will function just as automatically as any other organ. (p. 82).

As hypothesized by Krashen (1985), comprehensible input is essential to SLA. Comprehensible input means making the content understandable to students. Teachers may accomplish this in many ways. Echevarria, Vogt, and Short (2008) espouse three ways that teachers can make input comprehensible: use a rate of speech that is appropriate for their students' proficiency levels; provide a clear explanation of academic tasks; and, use a variety of techniques to make the content concepts clear for students. In the same fashion as Krashen (1985), they advocate the implementation of comprehensible input throughout each lesson taught to ELLs in order to help students understand the teacher and ultimately the content.

Rodrigo, Krashen, and Gribbons (2004) support the importance of comprehensible input. Their study involved college students taking Spanish as a foreign language in their fourth semester. One experimental group received an extensive reading approach that involved them reading as much as they could throughout the semester and meeting with the instructor to discuss their reading. Students in the group read a

combination of required texts and self-selected texts. A second experimental group used reading-discussion in which they read the assigned readings and took part in debates and discussions with their peers about the readings. The final comparison group took part in an intermediate grammar and composition course conducted along traditional teaching guides. Rodrigo et al. (2004) found that students in the two experimental groups outperformed students in the traditionally taught comparison group on the vocabulary and grammar tests administered to all groups of students at the end of the semester. Further, the students who took part in the reading-discussion experimental group outperformed the comparison group on the CLOZE test administered, yet the difference was not statistically significant. However, the group that received extensive reading did not differ on the CLOZE test when compared to the comparison group. Overall the researchers found that the experimental groups performed better than the comparison group in four of the six comparisons. They note that their results support the effectiveness of comprehensible-input based pedagogy in SLA at the intermediate level.

In his research, Krashen (1981) found that older learners tend to have distinct advantages when learning L2 as they typically have a relatively strong L1 foundation. Among the advantages he cited are that they: tend to learn more quickly, have greater knowledge of the world, can access more than one culture for advanced information, typically have greater control over the input they receive, and can more easily learn and apply the rules to aide in L2 acquisition. In like manner, Krashen, Long, & Scarcella (1979) found in their research that adults acquiring L2 progress through the early stages of morphological and syntactic development at a faster rate than L2 learners who are

children. They found the same to be true when comparing older children to younger children. Contrarily, they found when children are naturally exposed to a second language they achieve higher L2 proficiency in comparison to adults who begin acquiring L2 as adults.

Another significant theoretical model of SLA, also referred to as a framework, was proposed by Cummins (1981). He relates his theory to the work of theorists like Bruner (1975) and his communicative and analytic competence, and Donaldson (1978) and his embedded and disembedded language research. In his theory, Cummins (1981) posits that SLA involves two different types of proficiency. The first he categorizes as Basic Interpersonal Communicative Skills (BICS) and it is the language required to function in L2. It is the language needed for survival and for socialization. BICS is supported by interpersonal cues such as intonation present in face-to-face interaction, gestures, and facial expressions and is thus identified as context-embedded. He notes that L2 learners typically acquire BICS in English in about one to three years.

The second proficiency is what Cummins (1981) identifies as Cognitive Academic Language Proficiency (CALP). It is the language of academics that involves cognitive processes and high levels of thinking such as evaluation and synthesis. CALP is referred to as the language of school. According to Cummins (1981), it takes approximately five to seven years for ELLs to acquire CALP at a proficiency that is comparable to that of their native English-speaking peers. Unlike BICS, CALP is context-reduced. Subsequently, he notes that there is a sequential nature of the two with BICS developing first, and then CALP.

Cummins developed a continuum graph to represent academic and social language acquisition. On the horizontal axis of the graph is L2 acquisition which ranges on a continuum from context-embedded (BICS) to context-reduced (CALP). The vertical axis is a continuum ranging from cognitively undemanding (BICS) to cognitively demanding (CALP) language (Cummins, 1981). The graph demonstrates that language requiring BICS acquisition is found in quadrants A and B because it is cognitively undemanding and context-embedded. Language requiring CALP acquisition is located in quadrants C and D because it is cognitively demanding and context-reduced. Figure 1 is a representation of the continuum. Cummins (1981) included examples of specific types of language found in each quadrant.

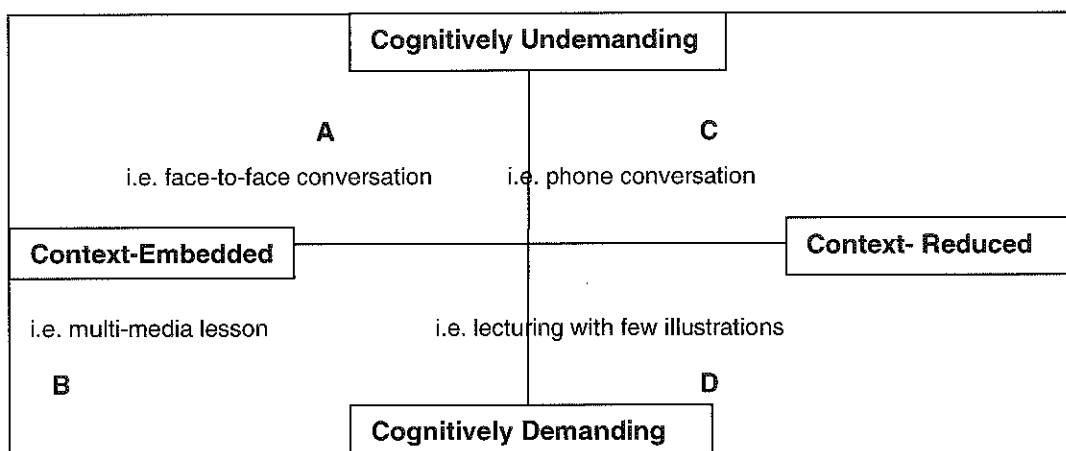


Figure 1. The dimensions of language proficiency.

In his study, Cummins (1981) analyzed data from 1,210 immigrants who were age six or younger when they arrived in Canada and who had not been exposed to English previously. He found that it took an average of five to seven years for the participants to develop CALP yet within two to three years the same students were fluent in

conversational English (BICS). His findings are supported by Biber (1986) who analyzed over one million words of English written text and speech that basically revealed the distinction between conversational and academic aspects of language proficiency. Cummins also relied on the work of Corson (1995) to support his findings. As per Cummins (1981), Corson (1995) noted the lexical differences between conversational interaction in English as compared to literary and academic uses of English. The various studies support Cummins' distinction between BICS and CALP acquisition.

Further supporting the theory by Cummins is a study by Hakuta, Butler, and Witt (2000) that included data from four different school districts, two in the San Francisco Bay Area in California and two in Canada. The districts were diverse in terms of their percentages of ELLs, percentages of students receiving free and reduced lunch, and ELL populations. In one of the districts in San Francisco, there were 1,872 students in the study ranging from first through sixth grade. The district had 35 percent of its students who received free and reduced lunch, and the majority of the ELLs were Vietnamese and Spanish speakers respectively. In the second school district in San Francisco, in which 74 percent of the students qualified for free and reduced lunch, 122 first, third, and fifth grade ELLs were included in the study. Both districts in California were considered among the most successful in teaching English to LEP students. A district in Ontario involved 285 ELLs who were mostly twelve years old, and the final district in Toronto included data from 1,200 immigrant students in grades five, seven, and nine. Although the districts were diverse, the findings among them were consistent. Hakuta et al.

(2000) found that in all four school districts, it took three to five years for ELLs to acquire oral English proficiency (BICS), and four to seven years for them to acquire academic English (CALP).

Aligned with Cummins' work, Collier (1987) conducted a study in a large, relatively affluent, suburban school district with an ESL program that was highly regarded. The study involved 1,548 students at the beginning of their L2 acquisition who received instruction in English in all subject classes. She gathered cross-sectional data from 1977 to 1986. The most common native languages spoken by the participants in the study were Korean, Vietnamese, and Spanish respectively. Collier (1987) considered the following variables in determining the time it takes ELLs to acquire academic English: age of arrival, English upon arrival, basic math skills in L1 upon arrival, and number of years of schooling in English (p. 617). Students' performance on the Science Research Associates tests in reading, language arts, math, science, and social studies administered in grades four, six, eight, and eleven were used to measure achievement. She found that students who arrived in the U.S. between the ages of eight and eleven required two to five years to reach the 50th percentile on national means in all subjects area tests.

Comparatively, students who arrived between the ages of five and seven were one to three years behind the performance levels of the ELLs who arrived between the ages of eight to eleven. Even more, children who were between twelve and fifteen years old when they arrived had the greatest difficulty and were projected to require as much as six to eight years to attain grade-level norms in academic achievement as their native English-speaking peers. Her research also indicated that overall it takes ELLs at least

four to eight years to reach national grade-level norms of native English speakers in all subjects of language and academic achievement.

Thomas and Collier (1997) conducted a study that took place from 1982 to 1996 involving ELLs from five large high schools. In their study, they defined success for ELLs as students reaching complete educational equivalence with their native English-speaking peers in all subjects. Subsequently, Thomas and Collier (1997) found that there are many factors that influence the amount of time it takes for L2 acquisition. The variables they identified are: proficiency in L1 and L2, age, socioeconomic status, and formal schooling in L1 (Collier, 1987; Thomas & Collier, 1997).

Further, Thomas and Collier (1997) assert that L2 acquisition is a complex developmental process and is only one of the processes taking place when a student transitions from one country to another. They point out that native English-speakers are not standing still waiting for ELLs to catch up with them; instead, they are developing cognitively and academically with every year of school, as well as continuing their acquisition of L1 in a learning environment that is favorable for instruction in English. With each passing school year, most native English speakers are developmentally expanding their language system. Standardized tests reflect this age-appropriate growth students are expected to make.

Critics of Cummins' contend that the distinction between conversational and academic language reflects a self-directed perspective on language that does not take into account its location in social practices (Edelsky, Altwerger, & Flores, 1991; Wiley, 1996). Opponents also reason that CALP more accurately represents what they refer to

as test-wiseness as it is more accurately a product of how students perform on content-area standardized tests (Edelsky et al., 1991). Finally, they assert that CALP represents a deficit theory because it attributes ELLs' academic failure to their low levels of CALP as opposed to inappropriate schooling (Edelsky, 1996; Edelsky et al., 1991).

There are many theories of SLA, some which are complimentary and others which are contradictory. Regardless of the theory districts, schools, and individual teachers choose to ascribe, it is clear that acquiring a second language is both a complex and a timely process.

Best Practices in ESL Education

Numerous researchers have investigated effective instructional practices for adolescent ELLs. Several best practices have been identified via research syntheses and reports.

Specifically, three major syntheses help identify salient pedagogical practices for instructing ELLs. Genesee (1999) edited a report compiled by a thirteen member panel which focused on oral language development, literacy development, and academic achievement of ELLs. The report was sponsored by the Center for Research on Education, Diversity and Excellence (CREDE), and addresses approaches for educating students from diverse linguistic and cultural backgrounds by presenting research findings and how to practically apply them in classroom settings (Genesee, 1999). A second synthesis was edited by August and Shanahan (2006) and it focuses on analyzing and synthesizing findings about English literacy attainment for ELLs. The study looked at the reading and writing skills ELLs need to be successful in school. A third synthesis was

compiled by Torgensen et al. (2007), and is a guidance document about *Academic Literacy Instruction for Adolescents*. In addition to providing instructional recommendations for improving academic literacy for forth through twelfth graders, the report makes specific commendations for instructing ELLs (Torgensen et al., 2007). The following are some of the most common best practices related to language acquisition suggested by various syntheses and reports.

Researchers assert that one salient pedagogical feature of effective ELL classrooms is that they provide clearly defined content and language objectives (Echevarria et al., 2008; Genesee, 1999; Genesee, Lindholm-Leary, Saunders, & Christian, 2006; Short & Echevarria, 2004). Teachers of ELLs must focus on the content students are expected to learn in the course, and they must focus on how to support students' language development. Largely, the objectives are intended to guide teaching as well as learning (Echevarria et al., 2008). While content objectives are not new to most teachers as they typically identify them for each unit and or lesson based upon their respective state and district standards, incorporating language objectives is not as common. Language objectives focus on language acquisition and more specifically, on the reading, writing, listening, and speaking students will demonstrate throughout the course of a lesson. By incorporating language strategies, teachers aid in students' English language development.

The significance of incorporating both content and language objectives is further supported by Echevarria, Short, and Vogt (2008). Their recommendation is based on a seven-year research study that was funded by CREDE (Short & Echevarria, 2004). An outgrowth of the research they conducted is a lesson planning and delivery approach known as the Sheltered Instruction Observation Protocol (SIOP) model (Short & Echevarria, 2004). SIOP consists of 30 features of instruction that are grouped into eight components found to be the most effective practices in sheltered instruction (Echevarria et al., 2008). The framework is intended to help teachers make information comprehensible for ELLs through the use of particular strategies and techniques. Specifically, Echevarria et al. (2008) posit that a prerequisite for attaining content standards is the acquisition of age appropriate knowledge of English. Therefore, teachers must intentionally plan for students' speaking practice as it is essential for development of academic English (Echevarria et al., 2008). In detail, the SIOP model stresses the importance of explicitly stating and posting student-friendly versions of the content and language objectives each day at the beginning of the lesson to set the stage for learning, and then revisiting the objectives at the end of the lesson to assess how well students progressed in achieving them (Echevarria et al., 2008).

In like manner, Francis, Rivera, Lesaux, Kieffer, and Rivera (2006) in their research-based recommendations for serving adolescent newcomers identify *Practical Guidelines for the Education of English Language Learners*. Aligned with incorporating language and content objectives, they recommend that classroom teachers of adolescent newcomers provide instruction that focuses on language and literacy development for content-area learning. They suggest that teachers identify one content objective and one language and literacy objective for each lesson they teach as this will help teachers focus their instruction on ELLs' academic and linguistic development. Furthermore, they recommend a content-based literacy approach to help students achieve success. In this type of approach "teachers evaluate their students' needs relative to content knowledge and to content-specific language and literacy demands" (Francis et al., 2006, p. 11). Another recommendation made by Francis et al. (2006) is that students receive instruction in academic language. By teachers incorporating academic language instruction, students are more likely to experience success in schools. Similarly, Torgensen et al. (2007), in their guidance document recommend that teachers focus on cognitive and linguistic skills in content area classes.

Overall, researchers agree that ELLs' academic success hinges on their mastery of academic language (Echevarria et al., 2008; Francis et al., 2006; Torgensen et al., 2007). They argue that students' acquisition and mastery of academic language is the most pivotal determinant of academic success for individual students (Echevarria et al., 2008; Francis et al., 2006; Torgensen et al., 2007). In order for ELLs to be successful academically, instruction must be designed to strategically promote oral language

development according to students' academic language needs (Echevarria, et al., 2008; Francis, et al., 2006; Genesee, et al., 2006; Torgensen, et al., 2007).

Another salient pedagogical feature of effective ELL instruction is an emphasis on vocabulary development (Echevarria et al., 2008; Francis et al., 2006; Genesee et al., 2006; Torgensen et al., 2007). Effective ELL teachers are expected to teach their students vocabulary along with content and language (Barcroft, 2012; Short & Echevarria, 2004). Barcroft (2012) argues for the necessity of giving vocabulary a central role when developing and designing programs, curriculum, and instruction in L2. The need for vocabulary to serve as a curricular anchor is further solidified by the relationship between vocabulary development and academic achievement (Short & Echevarria, 2004).

Vocabulary development and reading comprehension are related (Barcroft, 2012; Francis et al., 2006; National Reading Panel, 2000; Torgensen et al., 2007; Wallace, 2008). When students know more words it helps with their reading comprehension which ideally leads to students doing more reading which in turn leads to an increase in exposure to words (Torgensen et al., 2007). Essentially, having a strong vocabulary foundation is useful when students are learning to read (Wallace, 2008). An important factor to consider in the relationship between vocabulary and reading is that before native English-speakers begin learning to read, they are said to have acquired anywhere from 5,000 to 7,000 words on average, and as many as 10,000 words for those students on the higher end of the spectrum (Wallace, 2008). ELLs do not typically have the same basis in English. Since they do not have the same vocabulary word bank in English as native-English speakers, teachers must implement sound instructional practices to help

ELLs develop an adequate English vocabulary foundation before they can learn to read effectively (Bauer & Arazi, 2011).

When it comes to vocabulary instruction, researchers note that it must be: explicit, systematic, intensive, and extensive (Francis et al. 1006; Genesee et al., 2006; Torgensen et al., 2007). Teachers are tasked with the responsibility of carefully choosing which words to teach and how to teach them. Subsequently, Torgensen et al. (2007), in addition to several other researchers like Chung (2012) and Marzano and Pickering (2005), recommend that teachers invest the majority of vocabulary instruction time “teaching general-purpose academic words- such as analyze, frequent, and abstract- that are sophisticated in meaning but also appear in a variety of academic texts” (p. 13). These are typically referred to these as Tier One words (Chung, 2012; Marzano& Pickering, 2005). Moreover, teachers must decide which words to teach directly and which to teach using word-learning strategies such as prefixes and suffixes instruction (Francis et al., 2006).

Additionally, Barr, Eslami, and Joshi (2012) in their review of literature to determine if vocabulary instruction should be used to provide interventions for ELLs as a response to Texas English Language Proficiency Assessment System, identify several vocabulary instructional strategies. They note that vocabulary instruction should: link new vocabulary and background knowledge; focus on semantic relationships; restate dictionary definitions; use synonyms and antonyms; analyze the structure of new words; use contextual analysis; maintain personal content-related word lists; and allow students to work cooperatively to figure out words (Barr et al., 2012, p. 7-8). Likewise, Chung

(2012) advocates the necessity of complete vocabulary programs as opposed to teaching vocabulary in isolation. The salient components of effective vocabulary instruction that he identifies are: providing rich language word experiences, direct word teaching, instruction of word-learning strategies, and word-consciousness (p. 3).

Short and Echevarria (2004) further postulate that students should be given multiple opportunities to practice using their vocabulary words both orally and in print. Even more, researchers note that academic vocabulary which is challenging requires students to be exposed to the words and their meanings multiple times before they can develop a deep understanding of the words. Torgensen et al. (2007) report this number to be between twelve to fourteen times.

Overall, vocabulary instruction must be intentional, meaningful, and comprehensive. A delicate balance of teaching methods must be employed by teachers to help ensure that students make adequate gains in vocabulary development with each passing school year.

Another salient pedagogical feature of effective ELL classes is the use of scaffolding. Scaffolding is associated with Vygotsky's (1978) Zone of Proximal Development (ZPD). The ZPD is the difference between what a child can accomplish with help from someone more experienced and what the child can accomplish alone. Below the ZPD is students' level of actual development which characterizes the skills and knowledge students have already mastered. Above the ZPD is the level of potential development which includes skills and tasks students are not able to do independently, but can do with the guidance of an adult or in collaboration with peers. The ZPD is between

actual development and potential development and is the instructional level that produces the greatest learning gains. It allows students to build on their current knowledge to help advance their learning. Vygotsky (1978) notes that effective instruction should take place in the ZPD as it must be challenging, meaning that it is not too difficult or too easy. The goal is to help students achieve their level of potential development. One way this is accomplished is through the use of scaffolding.

Wood, Bruner, and Ross (1976) describe scaffolding as providing instructionally supportive activities to help guide students' effective learning and development in learning situations that are initially beyond their independent capacity. Echevarria et al. (2008) describe scaffolding as a situation where a teacher initially provides a significant amount of support to students as they begin a new concept or task, and as students' learning progresses, the teacher provides less and less support. As per Saye and Brush (2002), scaffolding has come to include teachers' use of guides, resources, tools, and other supplementary materials.

Walqui (2006), in her article that provides a conceptual framework for scaffolding instruction, describes three pedagogical scales of scaffolding. The three scales are identified as Scaffolding 1, 2, and 3. Scaffolding 1 is planned curriculum progression over time which is supposed to support the development of various skills. In Scaffolding 2, the activities in Scaffolding 1 are performed. It refers to the procedures used in specific activities. Scaffolding 3 is the assistance provided in the collaborative process of interaction to achieve Scaffolding 2. Walqui (2006) points out that the scaffolding does

not always move from one to three in a linear fashion; instead, how it progresses may fluctuate.

Pawan (2008) conducted a study to identify the chief types of scaffolding recognized by content-area teachers and to assess teachers' practical knowledge about scaffolding for ELLs. The study involved in-service teachers from seven different districts who were taking a required nine-month online course about ELL instruction. There were 33 content-area teachers involved: fifteen from elementary schools, six from middle schools, eleven from high schools; and one who taught adult education. Scaffolding was categorized as being linguistic (meaning it simplifies the English); conceptual (meaning it provides supportive frameworks for meaning); social (meaning it involves the engagement and support of others); or cultural (meaning the use of resources that are culturally and historically related to the learner) (Pawan, 2008, p. 1455). Pawan (2008) found that while approximately one-half of the teachers acknowledged that scaffolding was important for all students, only 19.2 percent indicated that all teachers were responsible for scaffolding. Additionally, Pawan (2008) found that teachers' knowledge of cultural scaffolding was surpassed by their knowledge of the other types of scaffolding.

Overall, scaffolding provides teachers with a means to integrate ELL instruction into content-area instruction, and it affords ELLs the opportunity to demonstrate their knowledge without completely relying on language (Pawan, 2008). It allows for the transfer of responsibility from the teacher to students as they progress toward their level of potential development.

The final feature of an effective ELL classroom that will be addressed in this review of literature is interaction. Merriam-Webster defines interaction as “mutual or reciprocal action or influence” (Interaction, n.d.). More specifically for ELLs, it is when students are given structured opportunities to practice academic discourse (Echevarria et al., 2008). Practice opportunities may occur between the teacher and students or among students (Echevarria et al., 2008).

Vygotsky (1978) cites interaction with peers as an effective way of developing skills and strategies. He suggests that teachers use cooperative learning exercises where less competent children develop with help from more skillful peers within the ZPD. The reason for this, he asserts, is because interaction allows students to progress from their actual development level to their potential developmental level (Vygotsky, 1978)

In Dale’s (1969) Cone of Experience model, he theorizes that learners retain 70 percent of what they say and write, and 90 percent of what they say and do using hands-on materials. He postulates that when learners are given opportunities to talk about, write about, and perform what they are learning, it leads to analysis, synthesis, and evaluation. Dale (1969) also notes that learners retain 10 percent of what they read, 20 percent of what they hear, 30 percent of what they see, and 50 percent of what they hear and see. The learner outcomes from reading, hearing, seeing, and hearing and seeing are typically knowledge, comprehension, and application. Interaction is a key component in learners retaining information.

Interaction is considered beneficial because it is recognized as facilitating adaptation and accommodation to individual differences in terms of abilities and styles.

It also allows for models and promotion of language for academic and cognitive purposes such as explanation, analysis, and formulation of complex ideas. Additionally, it extends ELLs' language use for academic and literacy purposes. August and Shanahan (2006) note that there is a significant relationship between reading and writing proficiency and oral proficiency.

Echevarria et al. (2008) identify the following benefits of interaction: brain stimulation, increased motivation, reduced risk (of feeling threatened when called upon to answer questions), more processing time, and increased attention. They note that it is advantageous to ELLs when they have opportunities to express themselves using English in quality academic situations (Echevarria et al., 2008). Additionally, their SIOP methodology promotes lessons that focus on high levels of interaction between teachers and students and among students.

Teachers are charged with creating opportunities for students' oral interaction and academic discourse (Short & Echevarria, 2004). In order to help students achieve academic English fluency, teachers must provide them with sufficient opportunities to practice using the language. This means that teachers must provide a balance between their own talk time and engaging students in academic talk during class time (Bennett, 2007; Short & Echevarria, 2004). Furthermore, the practice in academic discourse must be in ways that are meaningful to the students (Echevarria et al., 2008).

Agbatogun (2012) conducted a study that investigated the effectiveness of using the student response system in English language classrooms to improve interaction and student engagement. The student response system is a "wireless interactive handset that

collates and projects students' anonymous responses to a teacher's questions" (Agbatogun, 2012, p. 252) and then displays the captured data on a projection screen. It was an eleven week study involving 67 students in two Nigerian primary schools. The control group received instruction using the traditional lecture method while the experimental group received instruction that used the student response system which triggered interaction in the form of peer discussion. Each group participated in English Language Listening and English Language Speaking pre-tests and post-tests. There was a statistically significant difference between the pre-tests and post-tests of the control groups, yet no statistically significant difference between the pre-tests and post-tests of the control group. Additionally, the experimental group outperformed the control group on the post-tests. Agbatogun (2012) attributes the statistically significant differences to the interaction stimulated by the student response system.

In summary, Echevarria et al. (2008) state that, "through meaningful interaction, students can practice speaking and making themselves understood. That implies asking and answering questions, negotiating meaning, clarifying ideas, giving and justifying opinions, and more" (p. 121).

English-language learners are most successful when they take part in programs that are specifically designed to meet their needs (Genesee et al., 2006) both academically and linguistically. The program must be one that is "enriched, consistent, provides a challenging curriculum, and incorporates language development components and appropriate assessment approaches..." (Genesee et al., 2006, p. 377). Finally, all

research supports that, “the efficacy of each recommendation depends on its integration into a comprehensive program of instruction” (Torgensen et al., 2007, p. 94).

Blended Instruction

Blended instruction (BI), often used interchangeably with the term hybrid instruction, has many definitions. One of the most common is a combination of face-to-face and online instruction (Staker & Innosight, 2011; Watson, 2008; Watson, Murin, Vashaw, Gemin, & Rapp, 2010). Other definitions range from the more general combination of two or more delivery modes to more specific definitions which intimate the percentage of instructional time students spend online versus the percentage of instructional time students spend face-to-face with a teacher (Watson, 2008). Dziuban, Hartman, and Mokai (2004) assert that BI should be “viewed as a pedagogical approach that combines the effectiveness and socialization opportunities of the classroom with the technologically enhanced active learning possibilities of the online environment” (p. 5); thus further defining it is not necessary.

Although there is not a consensus on the definition of BI, there is agreement about its growth in K-12 education. Traditionally it has been predominant in higher education, yet within the last decade it has become increasingly pervasive in K-12 education (Horn, Staker, & Innosight, 2011; Means, Toyama, Murphy, Bakia, & Jones, K., 2010; Watson, 2008). According to the National Center for Education Statistics (2011), 45,000 students were enrolled in online courses in the year 2000 as compared to over one million in the year 2009. Further, in 2009, 45 out of 50 states had at least one form of online program available to its K-12 students (National Center, 2011). Several states such as Florida and

Idaho even require students to take an online course as part of their graduation requirements (Means, 2010). Even more, the growth in online learning has been impacted by state-led online programs like the Florida Virtual School that serviced 60,000 students in the 2007-2008 school year (National Clearinghouse of Statistics, 2013; Watson et al., 2010). Additionally, the increase in enrollment is partially attributed to the insurgence of BI (Horn et al., 2011).

Over the last ten years, there has been a significant growth in the use of BI (Barbour et al., 2011; Horn et al., 2011; Means et al., 2010). Some researchers attribute the increase to a combination of factors such as teacher shortages, reduced budgets, and even demands for results on high stakes tests (Horn et al., 2011). Others assert that BI is on the rise because a number of schools that could not afford the everyday expenses of operating a school needed an alternative instructional option which reduced or eliminated the daily expenses of running a school (Horn et al., 2011). For some institutions, that option is BI. Still other researchers characterize the increase as a response to the need for more innovate ways to meet the educational needs of the 21st century learner (Horn et al., 2011; Watson, 2008). Regardless of the reason however, its implementation has significantly increased.

Not only is BI growing in K-12 education in the United States, but it is on the rise internationally. The International North American Council for Online Education (iNACOL) conducted an international survey in 2011 to research how countries have implemented online and blending learning for K-12 students (Barbour et al., 2011). It was an update to the survey originally conducted in 2006 (Barbour et al., 2011).

Barbour et al., (2011) found that 35 out of 54 countries surveyed reported the availability of online and BI to at least some students which is an increase from the 2006 survey findings. Additionally, they found that internationally blended learning is occurring more frequently than online learning. They also report that even though blended and online instruction are found around the world, students from developed countries who live in urban areas are offered the most blended and online choices (Barbour et al., 2011).

While the implementation of BI is increasing internationally and nationally, research reveals that BI has been characterized in a variety of ways. For example, Watson (2008) notes that blending may occur at the course level, instructional level, or the level in which a student spends time in a traditional classroom versus online. He places it on a continuum and at one end is a fully online curriculum, and on the opposite end is a fully face-to-face curriculum. Table 1 shows the continuum (Watson, 2008, p. 12). Additionally, Horn et al. (2011) identify six models of blended learning that are shown in Table 2 (p. 11).

Regardless of the specific model of BI implemented, a host of advantages have been identified with using BI. The most notable is its ability to help educators personalize instruction and allow students to work at their own pace (Horn et al., 2011; Watson et al., 2010). It can provide teachers with specific, timely, individualized data about each student that may be used to individualize and differentiate instruction (Watson et al., 2010). Other advantages include BI's ability to: increase 21st century skills development, offer enriched experiences for the student, and increase communication and support (Watson et al., 2010). Further, Dziuban et al., (2004) note that BI provides:

A shift from lecture- to student-centered instruction in which students become active and interactive learners (this shift should apply to the entire course, including face-to-face contact sessions); it increases interaction between student-instruction, student-student, student-content, and student-outside resources, and integrates formative and summative assessment mechanisms for students and instructor. (p. 5)

Table 1
<i>Blended Instruction Continuum</i>
Fully Online Curriculum
<ul style="list-style-type: none"> Fully online curriculum with all learning done online and at a distance and no face-to-face component Fully online curriculum with options for face-to-face instruction, but not required Mostly or fully online curriculum with select days required in classroom or computer lab Mostly or fully online curriculum in computer lab or classroom where students meet every day Classroom instruction with significant, required online components that extend learning beyond the classroom and beyond the school day Classroom instruction integrating online resources, but limited or no requirements for students to be online Traditional face-to-face setting with few or no online resources or communication
Traditional Face-to-Face
Table 1

Table 2

<i>Models of Blended Instruction</i>	
Model	Description
Face-to-Face	Have face-to-face teachers to deliver most of their curricula. Teacher does online learning on a case-by-case study.
Rotation	Students rotate on a fixed schedule between learning online in a one-to-one environment, and sitting in a classroom with a traditional face-to-face teacher
Flex	Online platform that delivers most of the curricula and teachers provide online support on a flexible and adaptive as-needed basis
Online Lab	Relies on an online platform to deliver the entire course but in a brick-and-mortar lab environment
Self-Blend	Any time students take one or more courses online to supplement their traditional school's catalog; the online learning is always remote
Online Driver	Involves an online platform and teacher that delivers all curricula
Table 2	

As blended instruction is increasingly implemented in K-12 education, it is important to consider its effectiveness. In terms of online learning, most studies conducted over the years have found no statistically significant differences in effect between face-to-face and online learning (Means et al., 2010). Basically, learner outcomes for online instruction and face-to-face instruction were described as comparable in the studies (Means et al., 2010). Means et al., (2010) conducted a meta-analysis study in which they examined the effectiveness of online and blended online instruction in comparison to traditional face-to-face instruction. They reviewed studies that compared online learning and face-to-face learning, and blended learning with face-to-face learning. Forty-five studies met the criteria for inclusion in the meta-analysis. Their overall finding was that “online learning (the combination of studies purely online

and of blended learning) on average produces stronger student learning outcomes than learning solely through face-to-face instruction” (p. 24). They attribute the stronger outcomes largely to blended instruction. They also found that purely online learning has been equivalent to face-to-face instruction in effectiveness, and blended approaches have been more effective than instruction offered entirely in face-to-face mode. Furthermore, they found that a blended instruction had a larger advantage than strictly online and strictly face-to-face.

Staker and Innosight (2011) in their profiles of K-12 emerging BI models provide data about the effectiveness of various blended instructional models used around the United States. One school in particular, Carpe Diem Collegiate High School and Middle School(Carpe Diem), uses a blended model similar to the one in this study. It is an urban school located in Arizona serving grades six through twelve. Sixty one percent of the student population qualified for free and reduced lunch. In their BI model, students have a 55-minute period, rotating from online for concept introduction and instruction to face-to-face for reinforcement and application (2-3 rotations per day, 4 days of school per week, 8am-4pm). Although students receive a total of 1,007 hours of instruction, they attend school 145 days per school year. The school hires six full-time certified teachers to provide instruction in one of the following content areas: math, science, social studies, language arts, physical education, or electives. Each content teacher instructs all students (273) in his/her respective content across all grade levels on face-to-face instruction days. Furthermore, the school building has only five traditional classrooms and costs an estimated eight million dollars less than a traditional school being built in the same

neighborhood designed to accommodate approximately 200 more students than Carpe Diem.

In terms of the research regarding its effectiveness, Staker and Innosight (2011) report that Carpe Diem ranked first in its county in 2010 in student performance in math, and reading, and ranked among the top ten Arizona charter schools. This was similar to 2009 when, based on scores on the Arizona Instrument to Measure Standards (AIMS) test, Carpe Diem ranked first in the county in student performance for almost all grade levels and subjects. In almost all instances, at least 90 percent of the students at the school passed the tests in all grade levels and subjects tested. Both *BusinessWeek* and *U.S. News & World Report* recognized the school as one of the top high schools in 2009 and 2010 respectively.

Lim, Morris, and Kupritz (2007) conducted a study comparing online and blended learning looking specifically at instructional outcomes and student satisfaction. The study involved 125 undergraduate students at the University of Tennessee taking a program evaluation class. There were 59 students who took the course via an online delivery method and 69 students who took the course via a blended delivery method. At the end of the semester, researchers administered an online questionnaire that combined open-ended and close-ended items. They found that there was no statistically significant difference between the learners' perceived and actual learning based on delivery method. As per the survey results, both groups experienced a significant increase in their perceived and actual learning. Also, there was no statistically significant difference in the groups' perceived learning retention and learning application. However, statistically

significant differences between the groups were reported in relation to instructional difficulty, workload, and support. The students who experienced the online method reported a higher instructional difficulty level and a higher workload. Additionally, they felt less learning support during the course than students in the blended model. Overall, researchers concluded that the instructional delivery format did not affect students' learning or application to a significant degree (Lim, Morris, & Kupritz, 2007).

For a variety of reasons, BI is gaining popularity in K-12 education. While the definitions of BI vary, whether or not it significantly affects learning outcomes is debatable.

Trends indicate an increasing number of ELLs entering classrooms across the United States. Research suggests a growing trend toward blended instruction in K-12 education. The two are bound to intersect. Therefore, it is critical to determine how blended instruction impacts ELLs' language acquisition. The literature points out the eminent need for further investigation as there is a gap in research about blended instruction and English language learners in regards to language acquisition.

Chapter III

Methodology

Introduction

The purpose of this study is to investigate the effect of taking a blended English as a Second Language English (ESL) class or a blended ESL reading class on the language acquisition of English-language learners (ELLs) measured by the Texas English Language Proficiency Assessment System (TELPAS) in reading. Additionally, the study will investigate the effect of taking a blended ESL English class or a blended ESL reading class on ELLs' performance on the State of Texas Assessment of Academic Readiness (STAAR) in reading.

The overall design of the project was a quantitative approach, meaning that a formal, objective, systematic process where data are utilized to test hypotheses was implemented (Fraenkel & Wallen, 2008). Additionally, there was a qualitative component in the form of an exit interview conducted on the treatment group teacher that helps provide a clear understanding of the framework of the blended instruction classes.

Research Questions

This study investigates the relationship between blended instruction and the language acquisition of ELLs. Specifically, this study explores the following questions:

1. Is there a difference in the language acquisition of ELLs who take a blended ESL English class or a blended ESL reading class in comparison to ELLs who take a traditional ESL English or ESL reading class as measured by the Texas English Language Proficiency Assessment System reading exam?

2. Is there a difference in the language acquisition of ELLs who take a blended ESL English class or a blended ESL reading class in comparison to ELLs who take a traditional ESL English or ESL reading class as measured by the State of Texas Assessment of Academic Readiness English II reading exam?

Hypothesis

The research questions lay the framework for the subsequent null hypotheses:

Null Hypothesis One: There will be no statistically significant difference in the language acquisition of ELLs who took a blended ESL English or reading class as compared to ELLs who took a traditional ESL English or reading class as measured by the TELPAS reading exam.

Null Hypothesis Two: There will be no statistically significant difference in the language acquisition of ELLs who took a blended ESL English or reading class as compared to ELLs who took a traditional ESL English or reading class as measured by the State of Texas Assessment of Academic Readiness English II reading exam.

Sample Selection

The district. The district that is the subject of the study is a large urban school district in Texas. It officially became an independent school district in 1917 and was made up of one three-story brick school that was also used for church services, weddings, and community events. Over time, the district grew in conjunction with the community, and by the early 1970s it consisted of three elementary schools and one combination junior/senior high school. Although the district was originally located in a rural community on a sparsely populated flatland, urbanization eventually occurred in the late

1970s as more and more businesses materialized in the area, apartment complexes emerged, and city buses made the area more accessible. Census reports show that between 1970 and 1985 the population of the community quadrupled and continued to steadily grow over the years. In fact, the district's first Academic Excellence Index System (AEIS) report for the 1990-1991 school year reveals that there were 29,721 students in the district. Further, the same AEIS report shows the student demographics of the district at the time: 23.3 percent African-American, 19.1 percent Hispanic, 38.2 percent White, and 19.4 percent "Other." Additionally, 20.2 percent of the student population was labeled economically disadvantaged, 12 percent were Limited English Proficient (LEP), 11.5 percent were being served in bilingual/ESL classrooms, and 2.8 percent of the district's 1,829 teachers worked in bilingual/ESL classrooms.

The district has changed significantly since its inception almost 100 years ago. At the time of the study, the district had 41 campuses: 24 elementary, six intermediate, six middle, and five high schools (two with separate ninth grade centers) all located in a 36.6 mile radius. Further, the AEIS report for 2012-2013 reveals that the district served 45,783 students. Moreover, the report shows how the demographics of the district have also changed over time. In 2012-2013 the district's student population was: 31.2 percent African-American, 51.3 percent Hispanic/Latino, 3.5 percent White, and 13.5 percent "Other" who are now more specifically identified as .6 percent American Indian, 12.7 percent Asian, .2 percent Pacific Islander, and .6 percent Two or More Races. Additionally, 81.7 percent of the students were identified as economically disadvantaged,

36.2 percent were LEP, 36.0 percent were served in bilingual/ESL classrooms, and 18.3 percent of the district's 2,967 teachers worked in bilingual/ESL classrooms.

Taking a closer look at ELLs, the AEIS reports combine bilingual and ESL data as ELLs are either served in bilingual or ESL classes. In accordance with the subsection §89.1205 of Texas law, ELLs with 20 or more in any language classification in the same grade level district-wide from prekindergarten through sixth grade are served in bilingual classes (Texas Education Agency, 2012c). The law also states that for all ELLs the district is not required to offer bilingual education to, including those in secondary school, the district must provide an ESL program for (Texas Education Agency, 2012c). ESL classes are taught in English yet students receive linguistic support in the classes.

In terms of accountability, the district as a whole missed AYP in 2012 as per the then separated federal rating system. The reason for this is because the state imposes a ten percent cap based on the district's total student population on the number of special education students who may take the modified version of the standardized tests. Any modified tests administered beyond the 10 percent cap count as automatic failures on the exams, and the district exceeded the cap by two students. As a result, the district did not make adequate progress in the special education subgroup which resulted in missing AYP. There was no new state rating assigned in the spring of 2012 as the state suspended ratings for the 2012 performance data since the state was transitioning from Texas Assessment of Knowledge and Skills (TAKS) to STAAR. Instead, the state required the 2011 ratings to count for two years which means that the district was "recognized" in terms the state accountability system. A rating of recognized means that

the district's performance on TAKS ranged from 75 to 79 percent and that sufficient improvement was shown on TAKS from spring 2010 to spring 2011 to be at 80 percent in two years (Texas Education Agency, 2011a).

The high schools. The context of this study included three of the district's five high schools. Two of the district's high schools were not included in the study as they did not offer the classes that are the subject of the study; one of the schools is an early college high school and the other is a college preparatory high school. As for the three comprehensive high schools used in the study, the specific school each student attended was determined by a computerized lottery system. Consequently, all of the students in middle school are randomly selected to attend one of the three high schools during the spring of their eighth grade year, or more recently, in the fall of their seventh grade year. High School 1 (HS1) and High School 3 (HS3) are held on campuses that are comparable in size and they both also have separate ninth grade centers while High School 2 (HS2) houses ninth through twelfth graders and is the smallest of the three campuses. Therefore HS1 and HS3 receive more students in the lottery than HS2 because they are equipped to deal with larger numbers of students. An algorithm was created to assign students to one of the high schools with HS2 receiving 1.0 students for every 1.2 students in HS1 and HS3. The only way a student could go to a high school he/she was not selected to attend via the lottery was if he/she had a sibling who either graduated from one of the campuses or was currently attending one of the campuses. In such an event, a student could select to attend the school he/she was assigned via the lottery, or he/she could attend the school his/her sibling graduated from or was attending.

Even with the computerized lottery system the district uses for high school placement, the demographics of the three high schools were similar. Table 3 shows how their demographics compared to one another during the 2012-2013 school year.

Table 3								
<i>High School Demographics 2012-2013</i>								
High School	Total # of Students	African American	Asian	Hispanic/Latino	White	Eco Dis	LEP	LEP Served in ESL
1	3927	33.1%	12.9%	50.2%	3.0%	78.0%	11.7%	11.1%
2	2826	36.9%	12.1%	46.5%	3.4%	74.0%	9.5%	9.1%
3	4010	34.9%	12.2%	48.6%	3.5%	77.1%	11.4%	11.0%
Table 3								

Additionally, in terms of their numbers of ESL teachers: HS1 had 12 (out of 265 teachers), HS2 had 9 (out of 162 teachers), and HS3 had 12 (out of 229 teachers) (Texas Education Agency, 2013a).

When comparing federal accountability ratings for the schools, both HS1 and HS2 were rated as “met standard” for AYP based on their respective AEIS reports for the 2011-2012 school year. Although HS2 met standard for AYP for the 2011-2012 school year, it was considered a Stage 1, year two school because it missed AYP for the 2009-2010 and 2010-2011 school years in LEP math and LEP reading. As per TEA, if a school that participates in the federal Title I program misses AYP for two or more consecutive years for the same indicator (reading, mathematics, attendance, or graduation rate), it is identified for Title I School Improvement status (Texas Education Agency,

2011c). The school improvement categories range from Stage 1, the entry stage, to Stage 5, the final and most intensive stage of improvement. The required intervention activities increase at each subsequent stage as required by NCLB. TEA requires schools in Stage 1 to create a campus improvement plan that provides students the opportunity to transfer to another school in the district and to notify parents of the transfer, by October 27th of the school year (Texas Education Agency, 2011c). Stage 5 schools are required to undergo a major restructuring which includes options such as becoming a charter school and replacing most of the staff (Texas Education Agency, 2011c).

In order to be out of Stage 1, HS2 had to meet AYP for two consecutive years (2011-2012 and 2012-2013). Therefore, HS2 met AYP in 2011-2012, but was still a Stage 1, year two school at the start of the 2012-2013 school year. Additionally, HS3 did not meet AYP based on the 2011-2012 data. As this was their second consecutive year missing AYP, they became a Stage 1, year one school at the onset of the 2012-2013 school year. At the time of the study, HS1 was the only school out of the three that was not in school improvement.

The teachers. In the study, seven teachers' classes were examined. All of the teachers had their Texas ESL certification which means that they met the standard set by TEA in order to teach ESL classes. Additionally, they all earned "proficient" or "exceeds expectations" on their Professional Development and Appraisal System (PDAS) evaluations for the previous school year, 2011-2012. PDAS is the state's approved instrument for appraising teachers and identifying potential areas for professional development. Teachers are evaluated in eight domains: active, successful student

participation in the learning process; learner-centered instruction; evaluation and feedback on student progress; management of student discipline, instructional strategies, time/materials; professional communication; professional development; compliance with policies, operating procedures, and requirements; and improvement of all students' academic performance (Region 13 Education Center, 2013). Included in the appraisal are: a 45-minutes observation with documentation, walk-throughs, cumulative data, and scoring on all domains by the appraiser (Region 13 Education Center, 2013). Teachers receive one of four ratings defined in terms of their impact on students' learning: unsatisfactory, below expectations, proficient, and exceeds expectations (Region 13 Education Center, 2013). A rating of proficient means that the teacher exhibits behaviors that "result in considerable impact on students learning and which are demonstrated at a high percentage of time and with a high percentage of students (80-89%)" (Region 13 Education Center, 2013). As PDAS has been found to be both valid and reliable, it was the measure selected to help ensure that the teachers infor the study were comparable. This information was attained from the prospective teachers' respective administrators. Ideally by selecting teachers who had been rated as being at least "proficient" on PDAS, it helps address the extraneous variable of teacher quality.

Additionally, the teachers in the study were expected to incorporate the English Language Proficiency Standards (ELPS) as part of their curriculum. In an effort to help meet the academic needs of the growing ELL population and to close the achievement gap, the Texas State Board of Education approved the ELPS as a part of their Texas Essential Knowledge and Skills revised in 2007 (Texas Education Agency, 2007). As per

the ELPS, the expectation is that teachers in all content areas and programs provide linguistic accommodations that are commensurate with ELLs English language proficiency to help ensure that they are taught the academic English that is essential for academic achievement (Texas Education Agency, 2007). This translates into ELLs reading, writing, speaking, and listening in each class on a daily basis as part of the curriculum. Since the ELPS were a mandated and integral part of the curriculum, it was the expectation that all of the teachers in the study were implementing them as they were designed to be implemented.

In more detail, the teachers in the study received professional development in the Sheltered Instruction Observation Protocol (SIOP) methodology. The SIOP method is based on research by Echevarria, Short, and Vogt (2008). It is an instructional model for lesson design and delivery that ensures teachers use techniques to make instruction comprehensible for ELLs (Center for Applied Linguistics, 2013). The model allows ELLs to acquire academic knowledge while simultaneously developing proficiency in the English language (Echevarria et al., 2008). It emphasizes language development across the curriculum and provides students the opportunity to practice academic reading, writing, listening, and speaking skills. SIOP is comprised of thirty features across eight components: lesson preparation, building background, comprehensible input, interaction, strategies, lesson delivery, practice and application, and review and assessment (Echevarria et al., 2008). During a three-day professional development, teachers were taught each of the components and their respective features and how to implement them in content area classrooms. The district elected to implement SIOP because it is

researched-based and validated. Further, it has been shown to improve the performance of ELLs on measures of language and literacy (Echevarria, Ricardo-Tutor, Chinn, & Ratleff, 2011). Therefore, all of the teachers in the study took the initial three-day SIOP training and were expected to implement the methodology in their ESL classrooms.

Likewise, the ESL English and reading teachers in the study received training in the workshop model. The premise of workshop is that students do the reading, writing, and speaking for two-thirds of the class period, and the teacher uses direct instruction and modeling for one-third of the class period (Bennett, 2007). In the workshop model, the daily lesson involves: opening structure, mini lesson, work time, and debrief (Bennett, 2007). During the work time, teachers release students to work on the activity for the day (perhaps in small groups, pairs, or individually) while he/she goes around and confers with them. Then the teacher brings students back together in what is known as the “catch” to clarify, reteach or extend learning based on what was observed while conferring with students. After the “catch”, the teacher then “releases” the students to continue working. The model is designed to put more onus on students as one of its premises is that whoever is doing the reading, writing, listening, and speaking is doing the thinking, and learning is a consequence of thinking (Bennett, 2007). A consultant, Samantha Bennett, was hired by the district to help with the implementation of the model at the start of the 2009-2010 school year through the 2010-2011 school year. Teachers received training in the model by their respective English Language Arts/Reading (ELA/R) instructional specialists and via professional development provided by the district. It is the district’s expectation that all ELA/R teachers, including those who teach

ESL, implement the model. Fortunately, the SIOP method and the workshop model are not competing initiatives; instead, they are complementary.

On the whole, the teachers in the study were: ESL certified, rated as being at least proficient on their PDAS, expected to implement the ELPS, SIOP trained, and workshop model trained.

Traditional instruction model. There is not a district-wide lesson plan design teachers were expected to utilize. Instead, it was up to the discretion of each campuses' principal and leadership team exactly what lesson design was used at their respective campuses. Most of the lesson plan designs used at the high schools center around the Madeline Hunter lesson plan format or a variation of it. Regardless of the specific lesson plan design used at each campus, all ELA teachers, including those who taught English ESOL classes, were expected to incorporate the workshop model in their lesson plans. The teachers begin with an opening structure which is the same as Madeline Hunter's anticipatory set. In the opening structure/anticipatory set, teachers begin by preparing students for what they are going to learn by setting the purpose of the lesson and building the need to know the material. Next is the mini-lesson which is the same as Hunter's input/instructional process. Teachers instruct students during this time and model how to do the tasks for the worktime/guided practice. During the mini-lesson/input, teachers also check for understanding to ensure that students comprehend the tasks they are to complete. Worktime follows and is equivalent to Hunter's guided practice and monitoring. Students have an opportunity to implement what they learned during the mini-lesson/input and work either individually, in pairs, or in small groups (depending on

the assignment). Throughout worktime/guided practice, students are expected to read, write, and talk in an effort to help make meaning. While students are working, teachers confer with students to check for understanding, provide assistance as needed, and differentiate instruction. Teachers bring students back together as a class for short intervals during worktime/guided practice to either add to students' knowledge, provide clarification as needed, or to remediate. Finally, the lesson ends with a debrief which is the same as Hunter's closure. In the debrief/closure, students share their thinking, understanding, and/or task accomplished during worktime/guided practice. Teachers take the student work completed during worktime/guided practice and the knowledge gathered during the debrief as students share to help plan future teaching and learning. Specifically teachers note which students own the concepts and which students need more help.

Blended instruction model. In the blended classroom, the teacher still used the workshop platform: opening structure, mini lesson, worktime, and debrief. One difference is that during the opening structure and the mini lesson, the teacher used technology to introduce and model what students were to work on during the worktime. Also, during worktime students performed their tasks using the computer.

The treatment group teacher best describes the blended instructional model implemented in her classroom (H. Oliver, personal communication, October 30, 2013):

In the face-to-face portion [of the class], I modeled the objectives and outcomes. This was a highly interactive portion of the class. Students would talk to me and to each other. There would be a lot of questioning and answering. In the online

portion of the class, students managed their own learning, i.e. they controlled the pace of their learning and the tools that they used. They had a variety of tools to access [including] translators, language dictionaries, thesauri, image libraries as well as all the educational tools that Google offers. Students were required to engage in an online discussion or blog type writing and then respond to the posts of two peers. A minimum word limit of 200 words was set for the initial post and 100 words for a response post. Students were provided with sentence starters to use. I could join in the discussion and steer it in any direction. These discussions and blogs were used to write all our essays. I found that students were more willing to participate in the online discussion, as it gave them more time to organize their thoughts. The activities that we did in the online portion of the blended learning class focused on metacognition and critical thinking skills. This is the basis for successful learning. Metacognitive skills such as reflection, self-assessment and self-monitoring through blog activities helped the students analyze their own learning. I worked hard to create a safe environment where students could feel they could share the problems they were having with their own learning and then set personal, realistic goals. The discussion board was used to develop their critical thinking skills. The students posed questions to each other, and I would also ask questions, often playing devil's advocate, to ensure that the discussion was not superficial. In discussion posts students analyzed, sought information, made predictions, reasoned logically and transformed information. This technique was applied to the expository and the persuasive essays. If the

students finished the work early, they would work on customized grammar and writing exercises from Texas Writing Coach. At the beginning of the year, the students took a diagnostic test. The test then generated a customized program for each student. They then followed that program. As the teacher I could see how the students were progressing on their individual programs; which students were completing the tests successfully and which were not. This allowed me to have individual progress meetings with students.

Blended learning takes the best of face-to-face instruction and online instruction and combines it. Many people think that placing laptops in the classroom, or allowing students to play educational games is blended learning; they are mistaken. Blended learning is about tailoring instruction to the needs of the students. That is why I allowed my students to pick the technology tools so that they could find something that was compatible to their particular learning style or intelligence. Some students prefer to watch.

The students.For the purpose of this study, a convenience sample was used for the treatment group. However, the fact that students were randomly assigned to attend one of the three high schools helped provide some randomization. The treatment group was made up of LEP students who attended HS1 and were taking ESL English or reading class with Teacher A, who used the blended instructional approach. Students were assigned to take ESL English based on the following variables: TELPAS composite scores, TELPAS writing scores, *Idea* Proficiency Test-II scores, ESL instructional level, recommendations from current teachers, and/or final grades from their previous English

course (if applicable). Depending on the findings, students were placed in Beginning, Intermediate, or Advanced English for English Speakers of Other Languages (ESOL), or Sheltered/ESL English (E) for students who were considered Advanced High. Students who had been identified as Intermediate were typically placed in Intermediate ESOL English; those who had been identified as Advanced were typically placed in Advanced ESOL English; and those who had been identified as Advanced High were typically placed in Sheltered English. The students who were in the ESL reading class were assigned to the class because they failed the previous year's STAAR reading test and had been identified as students who would benefit from taking a reading class. Initial placement decisions were made by the ESL Department Chair at each campus, and were reviewed, modified as needed, approved, and finalized by the Language Proficiency Assessment Committee (LPAC) (Texas Education Agency, 2012a). The number of students in the treatment group broken down by the specific English or reading class they took is found in Table 4.

Since a convenience sample was used, a profile of the students was created using available demographic data in order to compare the treatment and comparison groups. Using a database, Cognos, the following data was collected: age, gender, ethnicity, home language, TELPAS years in school, whether or not the student was a refugee, and LEP status. The purpose of collecting the data and comparing it between the groups was so that the treatment and comparison groups could be closely matched on as many variables as possible. Ideally the researcher wanted to show that the treatment group and

comparison group were more alike than different on the variables that may impact the test scores being compared to test the hypotheses.

Table 4	
<i>High School I - Teacher A Treatment Group</i>	
Class	Number of Students
Intermediate English II ESOL	19
Advanced English II ESOL	18
English II –E	34
Reading II- E	20
Table 4	

The following data was collected for each member of the treatment group: fall 2012 ESL instructional level, fall 2013 TELPAS years in school, spring 2012 TELPAS reading scores, spring 2013 TELPAS reading scores, spring 2012 STAAR/EOC English I Reading scale scores, spring 2013 STAAR/EOC English II Reading scale scores, and home language.

Several steps were taken to identify the comparison group. First, using ELA/R master schedules from each high school, the same courses from the treatment group were identified at all high school campuses: Intermediate ESOL II (Inter. ESOL II), Advanced ESOL II (Adv. ESOL II), English II E (Eng. II E), and Reading II E (Read. II E). HS1 offered two sections (classes) of Inter. ESOL II and one section of Read. II E. HS2

offered one section of Inter. ESOL II, two sections of Adv. ESOL II, two sections of Eng. II E, and two sections of Read. II E. HS3 offered four sections of Inter. ESOL II, one section of Adv. ESOL II, and two sections of Eng. II E.

The ELA/R master schedule also denoted the teacher for each of the courses. The next step was to speak with the administrators over ELA/R at the three campuses and ask them to identify which teachers had received a rating of at least proficient on PDAS. Information from the administrators eliminated one reading teacher from HS1 thus leaving only the one section of Read II E available to be used in the comparison group. Further, the Read II E class had four students and was therefore eliminated as an option. Consequently, there were no available courses at HS1 that could be used as part of the comparison groups.

After identifying the final list of courses available to be used as comparison groups, the following variables were collected for each potential comparison group: LEP status, TELPAS years in school, home language, and ESL instructional level. The objective was to match the comparison group to the aforementioned variables of the treatment group. As the ELL population in the high schools was diverse, it was not possible to match the treatment group and the comparison group on all of the variables. From there one variable at a time was eliminated in an effort to match the groups on the remaining variables, but again the diversity of the ELL population prevented matching students on all of them. In the end, the only variable that was able to be matched was LEP status. As a result, a profile of the comparison groups was created in an effort to show that the treatment and comparison groups were comparable.

The final step in determining the comparison groups was to create a matrix with the courses, high schools, teachers, and number of sections. The objective was to choose at least one comparison group from HS2 for each course, and one comparison group from HS3 for each group. The specific breakdown of the comparison group is found in Table 5.

Table 5 Comparison Group				
Class	High School	Number of Students	High School	Number of Students
Intermediate English II ESOL	3 (E)	11	2 (T)	13
Advanced English II ESOL	3 (E)	11	2 (T)	9
English II –E	3 (E)	28	2 (T)	17
Reading II- E	2 (T)	14		
Table 5				

Therefore students who made up the comparison group were from HS2 and HS3 and were LEP students taking ESL English II or ESL reading classes with teachers who used the traditional instructional model. Students were placed in the classes using the same method as students in the treatment group. The criteria used for selecting the comparison group were students: (1) identified as LEP; (2) enrolled in one of the following ESL English classes: Intermediate English II for ESOL, Advanced English II for ESOL, English II-E, or Reading II-E, and (3) taking one of the aforementioned

classes with an ESL teacher who has received at least “proficient” on his/her 2011-2012 PDAS evaluation.

After all assessments were administered and scores attained, the following data was collected, analyzed, and compared for the treatment and comparison groups: spring 2012 TELPAS reading scores, spring 2013 TELPAS reading scores, spring 2012 9th grade STAAR/EOC scores, spring 2013 10th grade STAAR/EOC scores, and 2013 TELPAS years in school. All of the data was collected using Cognos reports and is archival in nature.

Instrumentation

The following instruments were used to gather data:

1. TELPAS Reading

During the 1999-2000 school year, the state implemented the TELPAS in reading to satisfy the federal requirement that ELLs be evaluated on their progress in becoming proficient in academic English (Educators Guide to TELPAS, 2012). TELPAS reading measures student performance in direct alignment with the English language acquisition skills and proficiency level descriptors defined by the ELPS, which are part of the Texas Essential Knowledge and Skills (TEKS) curriculum (Texas Education Agency, 2011b). The TELPAS reading assessment is designed to test students “ability to comprehend and interpret written text at grade-appropriate levels” (Texas Education Agency, 2011b, p. 12).

The TELPAS reading tests, which are multiple-choice, are designed to assess English language reading proficiency in a way that provides meaningful diagnostic

information about how well ELLs are learning to read the English they need for academic success (Texas Education Agency, 2011b). The test is built using four levels, or degrees, of linguistic accommodation, addressing the gradually reduced degree of linguistic accommodation that ELLs need as they progress from knowing little or no English to becoming fluent English readers. The English language proficiency levels reported are Beginning, Intermediate, Advanced, and Advanced High which are characterized by the degree of linguistic accommodation that students at that particular level need in order to read English with understanding (Texas Education Agency, 2011b). An explanation of each rating is found in Appendix A. The test blueprints require a certain number of test questions per proficiency level and per test objective. Score reports inform teachers of how successfully students demonstrate basic comprehension and analytical reading skills at the four proficiency levels.

Evidence supporting the validity of the TELPAS reading assessments has been collected by the state since the first administration of the tests, and continues to be collected with each passing year. The state reports that different types of validity evidence have been gathered regarding the test (Texas Education Agency, 2012d). Content validity of the TELPAS reading assessment is supported by its test design in that it provides staged linguistic accommodations commensurate with second language learning as it measures reading skills that students need for academic success in all subject areas. Additional evidence supporting content and construct validity comes from annual audits.

TELPAS reading reliability data is based on the internal consistency measure Kuder-Richardson Formula 20 (KR20)(Texas Education Agency, 2012d). Most internal consistency reliability measures are considered high if they are in the high .80s to low .90s range (1.0 being perfectly reliable)(Field, 2009). Kuder-Richardson Formula 20 for the TELPAS reading assessments range from .92 to .95 depending on the grade level of the test (Texas Education Agency, 2012d). Kuder-Richardson Formula 20 for the 2012 Reading TELPAS in grade nine was .944. The KR20 for grades ten through twelve test was .937(Texas Education Agency, 2012d). The Standard Error of Measure (SEM), which represents the amount of variance in a score resulting from factors other than achievement, is calculated using both the standard deviation and the reliability of test scores. Standard Error of Measure(SEM) assumes that underlying traits such as academic achievement cannot be measured precisely without a precise measuring instrument. The SEM for the 2012 TELPAS reading assessment was 2.846 and the p value was 73.611(Texas Education Agency, 2012d). The tests for grades ten through twelve had an SEM of 2.960 and a p value of 71.851(Texas Education Agency, 2012d). Kuder-Richardson Formula 20 and SEM show that the TELPAS reading tests in grades nine through twelve were both valid and reliable.

2. STAAR Reading English I (Grade 9) and English II (Grade 10)

Reliability for STAAR/EOC test scores was estimated using internal consistency, classical SEM, conditional SEM, and classification accuracy (Texas Education Agency, 2013d). The internal consistency estimates in the spring of 2012 at the reporting category level were lower than at the total score level thus indicating that interpretations of student

reporting category scores were not as reliable as at the total score level. As internal consistency estimates typically decrease as the number of test items decrease, internal consistency estimates made at the reporting category level can be noticeably different from those made at the level of the full assessment (Texas Education Agency, 2013d). Therefore, Texas Technical Committee Advisory (TTAC) noted that the lower reliability at the reporting category level must be taken into consideration when making interpretations of the scores (Texas Education Agency, 2013d). For the primary STAAR English assessments administered in spring 2012, the internal consistency estimates ranged from 0.81 to 0.93 which are in the good to excellent range (Texas Education Agency, 2013d). Internal consistency estimates across grades and content areas were found to be of a similarly high level, with no noticeable increases or decreases across grades or content areas (Texas Education Agency, 2013d).

Classical SEM was between 2 to 4 score points. The SEM values for the primary STAAR assessments (including STAAR L) administered in spring 2012 were 3.770 for the overall STAAR reading test (Texas Education Agency, 2013d). It is important to note that the SEM index provides only an estimate of the average test score error for all students regardless of their individual levels of proficiency (Texas Education Agency, 2013d). Additionally, Conditional SEM (CSEM) was reported to be in good range for the spring 2012 ninth and tenth grade reading assessments (Texas Education Agency, 2013d). CSEM provides a reliability estimate at each score point on a test (Texas Education Agency, 2013d). More specifically, CSEM is an estimate of the average test score measurement error that is conditional on the proficiency or scale score

estimate (Texas Education Agency, 2013d). Classification accuracy, which provides an estimate of the accuracy of students' classifications into performance categories based on current test results, was identified as being good (Texas Education Agency, 2013d). Furthermore, STAAR/EOC reliability data are based on the stratified coefficient alpha measure of internal consistency. The stratified coefficient alpha for STAAR/EOC tests involve a combination of dichotomous and polytomous (short-answer and extended response) items. Most internal consistency reliabilities were in the high .80s to low .90s range (1.0 being perfectly reliable) (Texas Education Agency, 2013d).

Annually, Texas collects validity evidence annually to support the various uses of STAAR test scores. TTAC provides ongoing input to TEA about STAAR validity evidence. In terms of validity, evidence was collected about the test content, response processes, internal structure, relationships with other variables, and analysis of the consequences of testing. All of the evidence gathered support that the STAAR/EOC exams are valid and reliable.

Data Analysis Procedures

Dependent Variables

The research analysis was conducted using ANOVA and differences between the treatment and comparison groups were examined. Further, ANCOVA was conducted with TELPAS 2012 and STAAR 9 as covariates in an effort to control for their effect on TELPAS 2013 and STAAR 10 scores. Variables have been identified as either dependent or independent. The dependent variables are the following: (a) TELPAS 2013

reading scores, (b) STAAR ninth grade reading scores, and (c) STAAR English II tenth grade reading scores.

Independent Variables

The following variables are categorized as independent (a) blended instructional method, (b) traditional instructional method, (c) gender, (d) age, (e) instructional level, (f) home language, (g) TELPAS years in U.S. schools, (h) refugee status, (i) STAAR English I reading scores, (j) STAAR English II reading scores. Furthermore, the variables were categorized as either scale or nominal.

Demographic Variables

Several demographic variables were collected from the experiment and comparison groups. Specifically, gender, age, instructional level, TELPAS years in U.S. schools, home language, and refugee status. Therefore, in addition to the dependent and independent variables, the demographic variables were analyzed in order to explore possible relationships, interactions, and differences across groups.

Chapter IV

Results

The purpose of the study was to compare the language acquisition of ELLs taking an ESL English or ESL reading class to the language acquisition of ELLs taking a traditional ESL English or ESL reading class. Language acquisition was measured using the TELPAS 2013 reading exam and the STAAR reading exam for English II.

Specifically, the following research questions were addressed:

1. Is there a difference in the language acquisition of ELLs who take a blended ESL English class or a blended ESL reading class in comparison to ELLs who take a traditional ESL English or ESL reading class as measured by the Texas English Language Proficiency Assessment System reading exam?
2. Is there a difference in the language acquisition of ELLs who take a blended ESL English class or a blended ESL reading class in comparison to ELLs who take a traditional ESL English or ESL reading class as measured by the State of Texas Assessment of Academic Readiness English II reading exams?

This chapter presents results using the following procedures which were described in Chapter Three: descriptive statistics, analysis of variance, and analysis of covariance with TELPAS 2012 and STAAR English I reading assessments as covariates.

Additionally, the following independent variables were analyzed to determine their effect on TELPAS 2013 and STAAR English II scores: gender, instructional level, age, and years in U.S. schools.

Descriptive Statistics Results

The descriptive statistics for the dependent and independent variables were analyzed. Appendix D contains the descriptive statistics for all participants, descriptive statistics for the group that received blended instruction (treatment group), and descriptive statistics for the group that received traditional instruction (comparison group).

Gender. In the study there were 84 females, 39 in the treatment group and 45 in the comparison group, and 110 males, 52 in the treatment group and 58 in the comparison group.

Home Language. A total of 17 different home languages were reported by the 194 students, including an “other” category which was used to describe the home languages of five students whose native languages were not delineated. Spanish was the most frequent home language reported with 96 native speakers, as indicated by the mode, followed by Vietnamese with 37 native speakers, and Burmese with 17 native speakers. In the treatment group, 40.7 percent of the students were native-Spanish speakers, compared to 57 percent of the students in the comparison group. Additionally, there was greater diversity in home languages within the treatment group with 16 out of the 17 languages represented in comparison to 11 out of the 17 in the comparison group. Home language frequencies are listed in Appendix E.

Instructional Level. Instructional levels indicate at which level of proficiency, beginning, intermediate, advanced, or advanced high, the students received classroom instruction. Students were assigned to various instructional levels based on their 2012

TELPAS composite scores, oral language proficiency test ratings, and recommendations, if applicable, of their previous year's teachers, or performance on Standard 10.

Descriptive statistics reveal that there were 20 participants at the beginning level, 40 at the intermediate level, 78 at the advanced level, and 56 at the advanced high level of instruction. Specifically, the treatment group consisted of four beginning, 16 intermediate, 43 advanced, and 26 advanced high level students in comparison to the comparison group that consisted of 16 beginning, 22 intermediate, 35 advanced, and 30 advanced high level students.

Age. Descriptive statistics indicate that the ages of the participants ranged from 14 to 19. Results show a unimodal frequency with mean, median, and mode being close in value and a SD of .929. The data were slightly positively skewed. Splitting the groups into those receiving blended instruction and those receiving traditional instruction showed the ages of the treatment group ranged from 14 to 18 with 43 percent of the students being 16. In the comparison group, the ages ranged from 14 to 19 with 41 percent of the students being 16. Since the comparison group had a student who was 19, the SD was higher at 1.00 than the SD of the treatment group at .825. Additionally, the mean age of the treatment group was 15.82 and the mean age of the comparison group was 16.01 indicating that the group that received traditional instruction was slightly higher than the group that received blended instruction which again was impacted by the comparison group having a student who was 19. Furthermore, data for both the treatment and the comparison groups were slightly positively skewed.

Refugee Status. There were 26 students identified as refugees, 12 in the comparison group and 14 in the treatment group. Since the data are nominal, the mode was the best measure of central tendency. The mode was zero (which was the category code for non-refugee) indicating that most of the participants were not identified as refugees.

Years in U.S. Schools. TELPAS years in U.S. school ranged from zero to six. TELPAS number of years in schools matched the number of years in U.S. schools prior to the 2012-2013 school year. For example, if a student was enrolled in a U.S. school for the first time during the 2012-2013 school year, the TELPAS years in school was coded as zero, and if a student entered a U.S. school during the 2011-2012 school year, TELPAS years in school was coded as one.

Descriptive statistics show that the mean and median for TELPAS years in U.S. schools were close in value, 3.38 and 3.30 respectively, while the mode was 6.00 with a SD of 1.87 indicating relatively large variations from the mean. The mode indicates that most students were either in their sixth year of school or greater. Skewness was relatively close to zero at .03. When the data was split based on instructional method received, the treatment group had a mean of 3.65, a median of 4.00, a mode of 6.00, and a SD of 1.78. The comparison group had a mean of 3.14, a median of 3.00, a mode of 2.00, and a SD of 1.93. Standard deviations for both groups were large indicating high dispersions of scores across the distributions. Six years in U.S. schools had the highest frequency in the treatment group at 20.9 percent, while in the comparison group it had a frequency at 20.4 percent second only to two years in U.S. schools at 25.2 percent.

STAAR English I Reading (Grade 9). Of the 194 participants, 145 had STAAR English I reading scores, 77 in the treatment group and 68 in the comparison group. There were 49 missing values, 14 in the treatment group and 35 in the comparison group. Missing data may be attributed to students not being enrolled in school at the time of the test, students not being enrolled in an English I class at the time of the test, students not having to take the STAAR English I reading but having to take the TAKS ELA assessment instead as part of their graduation requirement, or students being absent on the day of the test. The STAAR English I reading data reflects scores from the first administration of the exam in the spring of 2012.

Overall, the mean and median values were close, 1633 and 1631 respectively. The mode was 1750 and the standard deviation was 164.10. Skewness showed the data were slightly positively skewed. In the treatment group, the mean scale score was 1628.38, the median was 1631.00, and the mode was 1750.00 with a SD of 231.124 which indicates a large dispersion of scores across the distribution. For the comparison group, the mean and median were close in value, and it was bimodal with modes of 1490.00 and 1787.00, and skewness was slightly negative with a value of -.061. The SD was 172.67. It is worth noting that STAAR English I reading scores were indicative of how students performed prior to the beginning of the study (during the 2011-2012 school year).

STAAR English II Reading (Grade 10). Descriptive statistics show that 150 participants took the STAAR English II reading assessment, 76 in the treatment group and 74 in the comparison group. There were 44 missing values, 15 in the treatment group

and 29 in the comparison group. Overall, the mean and median were close in value at 1736 and 1722 respectively, and the mode was 1623. The data were slightly positively skewed. For the treatment group, the mean was 1731.32, the median was 1690.00, and the mode was 1623.00. In the comparison group, the mean was 1740.97, the median was 1775.00, and the mode was 1551.00. Both the treatment and comparison groups' data were slightly positively skewed and had high dispersions of scores with SDs of 204.55 and 221.01 respectively.

TELPAS 2012. There were 172 TELPAS 2012 scores reported, 84 in the treatment group and 88 in the comparison group. Twenty-two scores were missing, seven from the treatment group and 15 from the comparison group. Missing data may be attributed to students either not being enrolled in a Texas public school, or being absent during the TELPAS testing window. Overall, the results were close in value and unimodal with a mean of 2.96, median of 3.00, mode of 3.00, and SD of .975. In terms of the treatment group, the results for TELPAS 2012 rating showed a mean of 3.11, median of 3.00, mode of 3.00 and a SD of .836. For the comparison group, the mean and the median were close in value, 2.82 and 3.00 respectively, and the mode was 4.00 with a SD of 1.08. There was a slightly greater dispersion of scores in the comparison group than in the treatment group. The skewness values indicated that data for the treatment and comparison groups together and separately were slightly negative. As with STAAR English I reading scores, TELPAS 2012 scores represent how students performed prior to the start of the study.

TELPAS 2013. There were 190 TELPAS 2013 scores reported, 88 in the treatment group and 102 in the comparison group. A total of four scores were missing, three from the treatment group and one from the comparison group. Missing data may be attributed to students either withdrawing from school before TELPAS testing, or from students being absent during the TELPAS testing window. Overall the mean was 3.26, the median was 3.00, the mode was 4.00, and the SD was .818. In the treatment group, the mean was 3.36, the median was 3.00, and the mode was 4.00 with a SD of .681. For the comparison group, the mean was 3.17, the median was 3.00, and the mode was 4.00 with a SD of .913. Skewness for all sets of data was slightly negatively.

Analysis of Variance and Analysis of Covariance Results

To answer the research questions an analysis of variance (ANOVA) was conducted. Analysis of variance determines if means between the treatment group and the comparison group differed, and if so, whether the differences are statistically significant. Separate ANOVAs were run with TELPAS 2013 as the dependent variable and each of the following as independent variables: type of group (treatment or comparison), gender, instructional level, age, TELPAS years in U.S. schools, STAAR 9, STAAR 10, TELPAS 2012, and TELPAS 2013 (see Appendix F). Data run on the independent variables were analyzed in order to explore the relationships, interactions, and differences across groups.

TELPAS 2013

Research question one posed the question is there a difference in the language acquisition of ELLs who take a blended ESL English class or a blended ESL reading

class in comparison to ELLs who take a traditional ESL English or ESL reading class as measured by the TELPAS reading exam. The null hypothesis was that there is no statistically significant difference in TELPAS 2013 scores between the two groups.

According to Levene's Test of Equality of Error Variances (Levene's test), homogeneity of variance was violated, $p < .05$, which was an indication to reject the null hypothesis.

Since homogeneity of variance is an underlying assumption of all inferential tests, including ANOVA, any results on this question must be interpreted with caution.

ANOVA indicated that the mean scores for TELPAS 2013 were different between the treatment group ($M=3.31$) and the comparison group ($M=3.08$) signifying that in terms of actual scores, the treatment group scored higher than the comparison group on the assessment. Tests of Between-Subjects Effects indicate there was not a significant effect of instructional method on TELPAS 2013 scores, $F(1, 188) = 2.77$, $p = .098$, $r = .12$. The effect size was relatively small based on Cohen (1988) effect size scale meaning that approximately one percent of the total variance was explained by the instructional method. Table 6 shows the results of the ANOVA.

Table 6

Tests of Between-Subjects Effects TELPAS 2013 Type of Group
Dependent Variable: TELPAS 2013

Source	Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power ^b
Corrected Model	1.833 ^a	1	1.833	2.767	.098	.015	2.267	.380
Type of Group	1.833	1	1.833	2.767	.098	.098	2.767	.380
Error	124.530	188	.662					
Total	2143.000	190						
Corrected Total	126.363	189						

a. R Squared = .015 (Adjusted R Squared = .009); b. Computed using alpha=.05

Gender. There was homogeneity of variance as indicated by the Levene's test, $p > .0$. ANOVA indicated that there was a difference in TELPAS 2013 mean scores between females ($M=3.17$) and males ($M=3.33$). Although the actual scores indicated that males scored higher on TELPAS 2013, there was not a significant effect of gender on TELPAS 2013 scores, $F(1, 188) = 1.76$, $p = .186$, $r = .09$. Additionally, the effect size was small. Comparing the TELPAS 2013 scores of the treatment group to those of the comparison group by gender, actual scores for both males ($M=3.46$) and females ($M=3.24$) in the treatment group were higher than those of the males ($M=3.21$) and the females ($M=3.11$) in the comparison group (see Appendix G). The differences were not significant, $F(1, 186) = .269$, $p = .605$, $r = .03$.

Instructional Level. While the means among the groups were different, homogeneity of variance was violated, $p < .05$. In terms of actual scores, beginning students had a mean of 2.00, intermediate students had a mean of 2.85, advanced students had a mean of 3.43, and advanced high students had a mean of 3.75. The data indicated that the higher the instructional level of the students, the higher their TELPAS 2013 scores. There was a significant effect of instructional level on TELPAS 2013 scores, $F(3, 186) = 43.58$, $p = .000$, $r = .64$. The effect size was large as indicated by r . As a result of having a significant main effect, the Games-Howell post hoc test was run to explore and compare TELPAS 2013 means of all instructional level combinations. Games-Howell post hoc test indicated that differences in the means among students at the

beginning, intermediate, advanced, and advanced high instructional levels were all statistically significant with $p < .05$. The post hoc test further established that the higher the instructional level, the higher the TELPAS 2013 scores. Students at the advanced high instructional level had the highest scores. Table 7 provides the output from the Games-Howell post hoc test.

Table 7

Games-Howell Post Hoc Test Instructional Level/TELPAS 2013

Dependent Variable: TELPAS 2013

(I) Instructional Level	(J) Instructional Level	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Beginning	Intermediate	-.8500*	.18339	.000	-1.3382	-.3618
	Advanced	-1.4342*	.14907	.000	-4.8409	-1.0275
	High	-1.7455*	.15191	.000	-2.1584	-1.3325
Intermediate	Beginning	.8500*	.18339	.000	.3618	1.3382
	Advanced	-.5842*	.14411	.001	-.9646	-.2038
	High	-.8955*	.14705	.000	-1.2832	-.5077
Advanced	Beginning	1.4342*	.14907	.000	1.0275	1.8409
	Intermediate	.5842*	.14411	.001	.2038	.9646
	High	-.3112*	.10106	.013	-.5745	-.0480
High	Beginning	1.7456*	.15191	.000	1.3325	2.1584
	Intermediate	.8955*	.14705	.000	.5077	1.2832
	Advanced	.3112*	.10106	.013	.0480	.5745

*. The mean difference is significant at the .05 level

When comparing the treatment group to the comparison group, the actual scores for each proficiency level were higher for the treatment group than the comparison group (see Appendix G). Differences between the groups however were not significant, $F(3, 182) = .834, p = .477, r = .12$. The data show that in both the treatment and comparison groups, the higher the students' proficiency levels, the higher their actual scores on TELPAS 2013.

Age. The Levene's test specified that there was homogeneity of variance, $p > .05$. Further, the actual mean scores were different for each age group. There was a significant effect of age on TELPAS 2013 scores, $F(5, 184) = 3.32$, $p = .007$, $r = .29$. Since there was a main effect, post hoc tests were run to explore and compare TELAPS 2013 means among all age combinations. Based on the Tukey HSD post hoc test, the only means that were significantly different were between 15 year olds ($M = 3.51$) and 17 year olds ($M = 2.97$) with a $p < .05$ and a mean difference of .536. Comparing the treatment group to the comparison group showed that the actual scores were higher for the treatment group at ages 14, 15, 17, and 18 than the actual scores of the comparison group students of the same. However, 16 year olds in the comparison group had actual scores that were higher than their 16-year old counterparts in the treatment group (see Appendix G). It is important to note that the treatment group had only one 14 year old in comparison to the comparison group which had two. The differences however were not significant, $F(4, 179) = 2.42$, $p = .050$, $r = .23$.

Years in U.S. Schools. There was homogeneity of variance as indicated by the Levene's test, $p > .05$. Further, based on actual scores the means were different depending upon years in U.S. schools. However, there was no specific pattern for the differences in actual scores, i.e. as the number of years in U.S. school increased, TELPAS 2013 scores increased. Mean values are specified in Appendix G. The differences in TELPAS 2013 scores based on TELPAS years in U.S. schools were not significant, $F(6, 183) = 1.18$, $p = .321$, $r = .19$.

Comparisons between the treatment and comparison groups indicate that the students in the treatment group with zero, one, three, four, and five years in U.S. school had higher actual scores than students in the comparison group with comparable years in U.S. schools. However, the comparison group students had higher actual scores for two and five years in U.S. schools as compared to the treatment group. The differences were not significant, $F(6, 176) = 1.10$, $p = .367$, $r = .61$.

TELPAS 2012. Running the ANOVA showed that homogeneity of variance was violated, $p < .05$, based on the Levene's test. Additionally, the means were different among students at beginning ($M = 2.00$), intermediate ($M = 2.512$), advanced ($M = 3.28$), and advanced high ($M = 3.80$) ratings which indicates that in terms of actual scores, the higher students scored on TELPAS 2012, the higher they scored on TELPAS 2013. There was a significant effect of TELPAS 2012 score ratings on TELPAS 2013 score ratings, $F(3, 164) = 64.96$, $p = .000$, $r = .74$. An analysis of covariance was conducted with TELPAS 2012 as the covariate (see Table 8). The purpose was to control for the effect of TELPAS 2012 scores on TELPAS 2013 scores. ANCOVA indicated that TELPAS 2012 scores were significantly related to TELPAS 2013 scores, $F(1, 165) = 186.54$, $p = .000$, $r = .73$, with a large effect size. There was however not a significant difference in TELPAS 2013 scores between the treatment and comparison groups after controlling for TELPAS 2012 scores $F(1, 165) = .401$, $r = .04$.

When comparing TELPAS 2012 scores between the treatment and the comparison groups, actual scores indicate that the treatment groups' mean scores at the beginning ($M = 2.25$), intermediate ($M = 2.67$), and advanced ($M = 3.28$) levels were higher than the

comparison groups' means scores at the beginning ($M=1.93$), intermediate ($M=2.41$), and advanced ($M=3.27$) levels. However, the actual scores of students at the advanced high level were higher in the comparison group ($M=3.83$) than in the treatment group ($M=3.76$) (see Appendix G). The differences were not significant, $F(3, 50) = .826$, $p=.481$, $r=.69$.

Table 8.

ANCOVA TELPAS 2013 Covariate TELPAS 2012

Dependent Variable: TELPAS 2013

Source	Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power ^b
Corrected Model	60.406a	2	30.203	96.769	.000	.540	193.538	1.000
TELPAS2012	58.222	1	58.222	186.541	.000	.531	186.541	1.000
TypeofGroup	.125	1	.125	.401	.528	.002	.401	.096
Error	51.499	165	.312					
Total	1622.000	168						
Corrected Total	111.905	167						

a. R Squared = .540 (Adjusted R Squared = .534); b. Computed using alpha =

.05

STAAR English I (9) and STAAR English II (10). The Levene's test indicated homogeneity of variance, $p<.05$, when comparing means of TELPAS 2013 scores and STAAR English I reading scores, as well as when comparing means of TELPAS 2013 scores and STAAR English II reading scores. The actual scores for TELPAS 2013 and STAAR English I indicated differences in means as did the actual scores for TELPAS 2013 and STAAR English II. In terms of STAAR English I scores, there was a significant difference between STAAR English I reading scores and TELPAS 2013 scores, $F(35, 106) = 2.65$, $= .000$, $r=.68$. With regards to STAAR English II scores,

there was also a significant difference between STAAR English II scores and TELPAS 2013 scores, $F(35, 113) = 3.88, p = .000, r = .74$.

There was a great deal of variation in actual scale scores for the STAAR English I and STAAR English II reading exams. Based on actual scores, both the treatment and the comparison groups had mean scores that were higher and lower than one another depending on the STAAR English scale score. Both groups also had scale scores that were unique to either the treatment or the comparison group. The differences in mean scores for STAAR 9 were not significant, $F(19, 86) = 1.667, p = .058, r = .52$; neither were the differences in mean scores for STAAR 10, $F(21, 91) = 1.17, p = .301, r = .46$.

STAAR English II Reading (Grade 10)

The second research question asks if there is a difference in the language acquisition of ELLs who take a blended ESL English class or a blended ESL reading class in comparison to ELLs who take a traditional ESL English or ESL reading class as measured by the STAAR English II Reading exam. Null hypothesis two states that there is no statistically significant difference in STAAR English II reading scores between the groups. ANOVA with STAAR English II reading as the dependent variable indicated homogeneity of variance, $p > .05$, which dictated failing to reject the null hypothesis.

When comparing the means of the treatment group with those of the comparison group, the actual scores were different. Actual scores show the means for the comparison group ($M = 1741$) were higher than those of the treatment group ($M = 1731$). However, the differences were not significant, $F(1, 148) = .077, p = .782, r = .03$ (see Table 9).

Table 9.

Tests of Between-Subjects Effects Type of Group/STAAR 10
 Dependent Variable: STAAR 10

Source	Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power ^b
Corrected Model	3496.673 ^a	1	3496.673	.077	.782	.001	.077	.059
Type of Group	3496.673	1	3496.673	.077	.782	.001	.077	.059
Error	6706316.67	148	45312.948					
Total	458805878.0	150						
Corrected Total	6709813.040	149						

a. R Squared = .001 (Adjusted R Squared = -.006); b. Computed using alpha = .05

Gender. The Levene's test indicated homogeneity of variance, $p > .05$. Further, the actual scores showed the mean STAAR English II scores for males ($M=1742$) was higher than that of females ($M=1730$) indicating that males performed better on the exam than females based on actual scores. However, the difference was not significant, $F(1, 148) = .120$, $p = .695$, $r = .03$ (see Appendix H). As with overall STAAR English II scores for both groups, the effect size was small.

Comparing STAAR English II reading scores between the groups based on gender showed that based on actual scores, the mean for males ($M=1743$) in the treatment group was higher than the that of males ($M=1740$) in the comparison group. Conversely, actual scores for females in the comparison group yielded a higher mean ($M=1742$) than

that of females ($M=1717$) in the treatment group (see Appendix I). The differences were not significant, $F(1, 146) = .154$, $p=.695$, $r=.03$.

Instructional Level. The Levene's test indicated homogeneity of variance, $p>.05$. Based on instructional levels, the mean STAAR English II scores were different. Actual scores indicated that beginning level students had a mean of 1481, intermediate level students had a mean of 1609, advanced level students had a mean of 1719, and advanced high level students had a mean of 1864. Data show that as instructional level increased, so did the mean scores on STAAR English II reading. The means were significantly different, $F(3, 146) = 19.76$, $p=.000$, $r = .54$ (see Appendix H). A Tukey HSD post hoc test was conducted in order to explore and compare means for STAAR English II reading scores among students at the different instructional levels. Findings indicated that the differences among beginning, advanced and advanced high level students were significant, $p<.05$. Additionally, differences among intermediate, advanced, and advanced high students were significant. The difference between advanced and advanced high level students was also significant. Differences between beginning and intermediate level students were not significant with a $p > .05$. Table 10 provides the Tukey HSD post hoc results.

The actual scores indicated that the means for the treatment group were higher for students at the beginning, intermediate, and advanced high levels in comparison to the comparison group students at comparable levels. Conversely, the actual scores for the comparison group students at the advanced level provided a mean that was higher than

that of the students in the treatment group at the same level (see Appendix I). The differences were not significant, $F(3, 142) = .518$, $p = .670$, $r = .10$.

Table 10.

Tukey HSD Post Hoc Test STAAR10/ Instructional Levels All Participants
Dependent Variable: STAAR10

(I) Instructional Level	(J) Instructional Level	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Beginning	Intermediate	-127.520*	67.648	.239	-303.328	48.288
	Advanced	-237.713*	61.611	.001	-397.832	-77.594
	High	-382.5623*	62.334	.000	-544.559	-220.566
Intermediate	Beginning	127.520*	67.648	.239	-48.288	303.328
	Advanced	-110.193*	42.833	.053	-221.511	1.126
	High	-255.042*	43.866	.000	-369.044	-141.040
Advanced	Beginning	237.713*	61.611	.000	77.594	397.832
	Intermediate	110.193*	42.833	.053	-1.126	221.511
	High	-144.849*	33.822	.000	-232.749	-56.949
High	Beginning	382.562*	62.334	.000	220.566	544.559
	Intermediate	255.042*	43.866	.000	141.040	369.044
	Advanced	144.849*	33.822	.000	56.949	232.749

Based on observed means. *. The mean difference is significant at the .05 level.

Age. In regards to age, the Levene's test indicated homogeneity of variance, $p > .05$.

Further, actual scores noted differences in means based on age. The two 14 year olds in the study had the highest mean ($M=1909$) based on actual scores, and the one 18 year old had the lowest mean ($M=1600$). There was no specific pattern in mean scores, i.e. STAAR 10 scores were higher as age increased (see Appendix I). The mean differences among the age groups were not significant, $F(4, 145) = .928$, $p = .449$, $r = .16$ (see Appendix H).

Based on actual scores, means were higher in the treatment group for students at ages 14, 15, and 17 than students in the comparison group at the same ages. Contrarily, 16 year olds in the comparison group had a higher mean based on actual scores in comparison to 16 year olds in the treatment group (see Appendix I). The differences were not significant however, $F(3, 141) = 1.676$, $p = .175$, $r = .18$.

TELPAS Years in U.S. Schools. The Levene's test indicated homogeneity of variance, $p > .05$. Additionally, based on actual scores, means for STAAR English II reading were all different based on number of years in U.S. schools (see Appendix I). In terms of actual scores, the highest mean reported was for students with zero years in U.S. schools ($M = 2055$), and the lowest mean reported was for students with three years in U.S. schools ($M = 1636$). It is important to note that there were only three students with zero years in school in comparison to 29 with three years in school. Differences in mean based on TELPAS years in U. S. schools were significant, $F(6, 143) = 2.60$, $p = .020$, $r = .31$ (see Appendix H). Since the main analysis denoted that the differences in mean were significant, the Tukey HSD post hoc test was conducted in order to explore and compare means on STAAR English II reading based on the different TELPAS years in U.S. schools. Based on Tukey HSD, the only difference that was significant was between students with zero years in U.S. schools and those with three years in U.S. schools with a mean difference of 419 as seen in Appendix J.

Actual scores for students in the treatment group showed higher means for STAAR English II reading scores for students with one, two, and five years in U.S. schools as compared to students in the comparison group with comparable years in U.S.

schools (see Appendix I). Conversely, students in the comparison group with zero, three, four, and six years in U.S. schools had higher means in terms of actual scores in comparison to students in the treatment group with the same years in U.S. schools. The differences were not significant, $F(6, 136) = .885$, $p = .507$, $r = .19$. It is important to note that in the experiment there were two students with zero years in school in comparison to the comparison group that had one student with zero years in school.

TELPAS 2012/TELPAS 2013. The Levene's test indicated homogeneity of variance, $p > .05$, when comparing means for STAAR English II reading scores and TELPAS 2012 scores, and when comparing means for STAAR English II reading scores and TELPAS 2013 scores. Actual scores for STAAR English II reading provided means that were different based on students' TELPAS 2012 scores and their TELPAS 2013 scores. In both cases, the higher the students' TELPAS reading score ratings, the higher their STAAR English II reading scores as shown in Appendix I.

The differences in STAAR English II reading means based on TELPAS 2012 score ratings of beginning, intermediate, advanced, or advanced high, were significant, $F(3, 135) = 40.41$, $p = .000$, $r = .69$ (see Appendix H). Therefore the Tukey HSD post hoc test was conducted to explore and compare STAAR English II reading means based on the different TELPAS 2012 scores. Findings indicated that differences in means on STAAR English II reading were significant among students who scored at the beginning level, and those who scored at the advanced and advanced high levels. Differences were also significant among students who scored at the intermediate level and those who scored at the advanced and advanced high levels. Lastly, there was a significant

difference between students who scored at the advanced level and those who scored at the advanced high level. Furthermore, the actual scores of the treatment group indicated higher means than those of the comparison group. However, differences in STAAR English II reading scores based on TELPAS 2012 scores between the treatment and comparison groups were not significant, $F(3, 131) = 19.492$, $p = .845$, $r = .08$.

With regards to STAAR English II reading score means based on TELPAS 2013 scores, the differences were significant, $F(3, 145) = 36.45$, $p = .000$, $r = .66$. The Tukey HSD post hoc test was performed to compare the means. Results indicated that the differences in the STAAR English II reading test means were significant among students who scored at the advanced high level of TELPAS and those who scored at the beginning, intermediate, and advanced levels of TELPAS, $p < .05$ (see Table 11). No other differences were significant. Comparing the treatment group to the comparison group indicated differences in means based on actual scores (see Appendix I). The differences however were not significant, $F(2, 142) = 2.637$, $p = .075$, $r = .19$.

Table 11.

Tukey HSD Post Hoc Test STAAR 10/TELPAS 2013
Dependent Variable: STAAR10

(I) TELPAS 2013 Rating	(J) TELPAS 2013 Rating	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Beginning	Intermediate	-97.667	117.604	.841	-403.849	208.516
	Advanced	-195.255	114.752	.327	-493.504	102.994
	High	-425.240*	114.057	.002	-721.683	-128.797
Intermediate	Beginner	97.667	117.604	.841	-308.516	403.849
	Advanced	-97.588	41.276	.089	-204.867	9.690
	High	-327.573*	39.302	.000	-429.723	-225.424
Advanced	Beginner	195.255*	114.752	.327	-102.994	483.504
	Intermediate	97.588*	41.276	.089	-9.690	204.867
	High	-229.951*	28.893	.000	-305.080	-154.890

High	Beginner	425.240*	114.05743	.002	128.797	721.683
	Intermediate	327.573*	39.30233	.000	225.424	429.723
	Advanced	229.951*	28.893	.000	154.890	305.080

Based on observed means. *. The mean difference is significant at the .05 level.

STAAR English I Reading (9). The Levene's test indicated that homogeneity of variance was violated, $p < .05$. Additionally, based on actual scores, the means for STAAR English II reading were different based on STAAR English I reading scores. Differences in means were significant, $F(33, 103) = 4.05$, $p = .000$, $r = .74$.

An analysis of covariance (ANCOVA) was conducted with STAAR English I as the covariate. The purpose was to control for the effect of STAAR English I reading scores on STAAR English II reading scores. The Levene's test indicated homogeneity of variance, $p > .05$. ANCOVA indicated that STAAR English I scores were significantly related to STAAR English II scores, $F(1, 134) = 89.18$, $p = .000$, $r = .63$ as indicated in Table 12. There was however not a significant difference in STAAR English II reading scores between the treatment and comparison groups after controlling for STAAR English I reading scores $F(1, 134) = .329$, $p = .567$, $r = .04$.

Table 12.

ANCOVA STAAR 10 Covariate STAAR 9
Dependent Variable: STAAR 10

Source	Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power ^b
Corrected Model	2402688.34 ^a	2	1201344.168	44.607	.000	.400	89.213	1.000
STAAR9	2401846.308	1	2401846.308	89.182	.000	.400	89.182	1.000
Type of Group	8860.621	1	8860.621	.329	.567	.002	.329	.088
Error	3608884.745	134	26931.976					
Total	414687129.0	137						
Corrected Total	6011573.080	136						

- a. R Squared = .400 (Adjusted R Squared = .391); b. Computed using alpha = .05

Comparing the treatment and comparison groups indicated that based on actual scores, both groups had means that were higher and lower than one another depending on the particular STAAR English II reading scale score. Both groups also had scales scores that were unique to either the treatment or comparison group. However, the differences were not significant, $F(19, 83) = 1.679$, $p = .057$, $r = .53$.

Summary

The study explored the differences in language acquisition of students who received a blended method of instruction in ESL English or ESL reading as compared to those who received instruction via the traditional method based on the TELPAS 2013 and STAAR English II reading exams. Results indicate that there were no statistically significant differences in TELPAS 2013 scores between the group that received the blended method of instruction and the group that received the traditional method of instruction. The same was true in terms of STAAR English II reading scores, the differences between the groups were not statistically significant. ANOVA and ANCOVA did indicate a significant effect of TELPAS 2012 scores on TELPAS 2013 scores. Likewise there was a significant effect of STAAR English I reading scores on STAAR English II reading scores.

Further, results indicated that there were significant differences in TELPAS 2013 reading scores based on instructional level, age, STAAR English I reading scores, and STAAR English II reading scores, but there were not significant differences based on

gender and years in U.S. schools. As for STAAR English II reading scores, there were significant differences in scores based on instructional level, years in U.S. schools, and TELPAS 2012, yet no significant differences based on gender and age.

Chapter V

Conclusions, Interpretations, Recommendations

Introduction

The purpose of this study was to investigate the effect of taking a blended English as a Second Language(ESL) English class or a blended ESL reading class on the language acquisition of English-language learners (ELLs). Specifically, the research questions were: is there a difference in the language acquisition of ELLs who take a blended ESL English class or a blended ESL reading class in comparison to ELLs who take a traditional ESL English or ESL reading class as measured by the Texas English Language Proficiency Assessment System (TELPAS) reading exam; and, is there a difference in the language acquisition of ELLs who take a blended ESL English class or a blended ESL reading class in comparison to ELLs who take a traditional ESL English or ESL reading class as measured by the State of Texas Assessment of Academic Readiness Reading (STAAR) English II reading exam?

To answer research question one, scores from TELPAS 2013 were used since all ELLs in the state of Texas are required to participate in the exam annually to measure their language acquisition. Further, the STAAR English II reading exam scores were used to answer research question two because the assessment is one of the end-of-course exams most of the students in the study were required to take and pass as part of their graduation requirements. Furthermore, the STAAR assesses students “readiness for success in subsequent grades and courses, and ultimately, for college and career” (Texas Education Agency, 2010).

Language Acquisition as Measured by TELPAS 2013 Interpretations

Research question one inquired is there a difference in the language acquisition of ELLs who take a blended ESL English class or a blended ESL reading class in comparison to ELLs who take a traditional ESL English or ESL reading class as measured by the TELPAS reading exam. The analysis of variance indicated that although based on actual scores the mean of the group that received the blended instructional method ($M=3.36$) was higher than the mean of the group that received the traditional instructional method ($M=3.17$), the difference was not statistically significant. These findings are incongruous with the findings of Means, Toyana, Murphy, Bakia, and Flores (2010) who determined in their meta-analysis that blended approaches were slightly more effective than instruction offered in the face-to-face, traditional mode as noted in the review of the literature.

One consideration is that the delivery method may not be significant. Research indicates that students attain English in a sequential manner when they are linguistically and developmentally ready. Specifically, Krashen's (1985) natural order hypothesis asserts that the rules of language are acquired in a predictable order in which some rules are acquired early in L2 development and others are acquired later. Additionally, Cummins (1981) established that language acquisition is sequential in that students acquire Basic Interpersonal Communication Skills (BICS) first and then Cognitive Academic Language Proficiency (CALP). Therefore because of the sequential nature of language development, it is possible that regardless of the instructional method, L2

learners acquire language only when they are linguistically and developmentally ready to do so.

Another consideration is that the teachers in the study may have implemented best instructional practices for ELLs. Research indicates that when teachers implement salient pedagogical practices, such as incorporating content and language objectives (Genesee, 1999), vocabulary development (Torgensen et al, 2007), scaffolding (Vygotsky, 1978), and interaction (Vygotsky, 1978), students acquire academic content and language successfully. Such practices may be implemented in classes in which the instructional delivery is blended or traditional. It is likely that the implementation of best practices was more significant than differences in instructional delivery method.

Additionally, the district has in place other instructional frameworks, specifically Sheltered Instruction Observation Protocol (SIOP), which has been endorsed as an effective methodology for helping ELLs acquire academic English effectively. Studies have shown that students of teachers who implement SIOP the way it was intended and with high frequency yielded higher scores on curriculum-based tests in comparison to those students whose teachers either implemented the model with low frequency and/or fidelity or not at all (Echevarria, Richard-Tutor, Chinn, Tattleff, 2011). Similar gains were reported on standardized tests (Echevarria et al., 2011). All teachers in the study received SIOP training. Conceivably, the teachers in the study implemented the SIOP methodology with similar fidelity and frequency which could mean that all students in the study performed similarly because of SIOP implementation. As with the best

practices, perhaps the implementation of the SIOP methodology offset the effect of instructional delivery method.

The relatively small size of the sample is another consideration when taking into account the lack of statistically significant difference between the treatment and comparison groups. There were 91 participants who received blended instruction and of those 91, 88 had TELPAS 2013 reading scores. Additionally, there were 103 participants who received traditional instruction and 102 of those had TELPAS 2013 reading scores. Small samples produce greater sampling variability and often produce small effect sizes. It is possible that the small n values prohibited a statistically significant difference from being observed. A larger sample may have produced statistically significant differences if indeed they existed.

Effectiveness of the teachers in the study is another consideration. While PDAS was used to ensure that the teachers received a rating of at least proficient, there were three different administrators who assessed the teachers in the study. This means that although the measure is valid and reliable, different administrators may have slightly different interpretations of the PDAS rubric. Additionally, even if all of the teachers in the study had had the same appraiser, there is a wide range of variability within the PDAS rating of proficient. While some teachers may have been in the upper echelon of the proficient scale, others may have been in the middle, while yet others may have been in the bottom, yet all of the teachers were rated as proficient. Therefore, variability among teachers may account for the lack of statistically significant differences.

The absence of a true experimental research design is another consideration when analyzing the lack of statistically significant difference between the groups. Neither students nor teachers were randomly assigned to be part of an experimental or control group. Although there was some randomization because of the way the district assigns students to high schools, a convenience sample was used as the treatment group. Furthermore, it was not possible to match students in the treatment and comparison groups on more than two variables because of their diversity.

A final consideration is that the data reflects the first year of implementation of the blended instructional model. The teacher who implemented the model created the class on her own. Although she followed the district's pacing guides, she did independent research to determine what activities to include in the online portion of the class. Her objective was to find activities that were research-based and research-validated. Building the blended course required her to work extensive hours outside of the normal work day. Since the study reflects the first year of implementation, the teacher did not have the opportunity to refine the course over time. Perhaps in subsequent years if she would have continued implementing blended instruction and refining the course, statistically significant differences may have been observed.

TELPAS 2013 and the Independent Variables.

Gender. Actual scores indicated that when considering all participants, males performed better on the TELPAS 2013 reading assessment, yet the difference was not significant. The same was true when comparing scores by gender between the group that

received blended instruction and the group that received traditional instruction.

Statistically significant differences in scores based on gender were not expected.

Age. The differences in TELPAS 2013 mean scores based on age were significant, specifically between 15 and 17 year olds with a mean difference of .536. Thomas and Collier (1997) identified age as a contributing factor in the language acquisition of ELLs. They found that it takes older students longer to acquire English, especially if they come to the U.S. when they are older, which may account for the differences between the 15 and 17 year olds. Specifically, when students begin to learn English when they arrive in the U.S. between the ages of 12 and 15, they were projected to need six to eight years to attain academic English in a way that was commensurate with their native-English speaking peers. Older students often complete high school before acquiring academic English in a way that is comparable to their peers who are native-English speakers.

Differences in mean may also be attributed to apathy. Sometimes as students get older and they are not at the same academic level as their peers, they become apathetic about school. A few students even bide their time until they turn 18 so that they can drop out of school. This may be a contributing factor for the disparity in mean scores between 15 and 17 year olds.

TELPAS Years in U.S. Schools. It is important to note that as per the Texas Education Agency (TEA), if a student was enrolled for at least one instructional day during a school year, it counted as one full school year. Therefore, a student who enrolled the last week of school and a student who enrolled the first week of school

during the same school year were both identified as students in their second year in U.S. schools at the start of the subsequent school year. This is important when considering the results because while a TELPAS year in U.S. schools for some students may reflect an entire academic school year, for other students it may reflect months, weeks, or even days in school. Furthermore, the maximum number of TELPAS years in U.S. schools that can be reported is six. Therefore, if a student had been in U.S. schools beyond six years, the student was coded with a six for TELPAS years in U.S. schools. While six TELPAS years in U.S. schools may reflect six academic calendar school years, it may also reflect anywhere from seven to ten academic calendar school years.

The data indicated that there was an effect of TELPAS years in U.S. schools with students who had been in school for three years performing best on the TELPAS 2013 exam. Cummins (1981) noted that BICS develops more rapidly than CALP. Students typically require one to three years to develop BICS and five to seven years to develop CALP. Students within their first three years of school are acquiring English more rapidly than those who have been in U.S. schools longer (Cummins, 1981). The reason for this is because those students who have been in U.S. schools beyond three years have typically already acquired BICS and are more often acquiring CALP which is more academically demanding. However at year three, most students are at the height of BICS acquisition which could account for the high means on TELPAS reading.

Instructional Level. Data indicated that the higher the students' instructional levels, the better they performed on TELPAS 2013 reading. This was true when considering all participants together, as well as when considering them in terms of

instructional method received. It stands to reason that instructional level had an effect on TELPAS 2013 reading scores. Students were placed in their respective ESL English or ESL reading classes based on how much English they had acquired prior to the start of the 2012-2013 school year. This indicates that the teachers of these students taught their respective content at a level commiserate with the instructional level of the course (i.e. Intermediate English II ESOL or Advanced English II ESOL). At the very least, students were expected to be able to read, write, speak, and listen in accordance with the instructional level of the class. Therefore it is reasonable that the higher the students' instructional levels, the higher the students' ratings on TELPAS 2013 reading

Home Language and Refugee Status. Initially the plan was to run an ANOVA using home language as an independent variable, but because there were 16 languages spoken and one "other" category within the relatively small sample population, the data was not run. Standard error would have been large indicating that the data would not have provided an accurate reflection of the population from which the sample came. In other words, there were not enough participants in each language category to provide data representative of the population.

As for refugee status, there were too few students identified as refugees (12 in the treatment group and 14 in the comparison group) meaning that the sample size was too small to be stable. Additionally, the effect size would have been too small to glean any real outcome. It is worth noting that information about whether or not the refugee students had any previous schooling was not collected as the data were not available in the database. Prior schooling is significant when it comes to language acquisition.

Thomas and Collier (1997) found that students who have been educated in their native languages and are capable of reading and writing in the languages typically acquire English more rapidly than students without schooling in their native languages. Many of the refugee students in the district were born and raised in refugee camps, some with organized education and others that did not have organized education. Therefore, refugee status, and more specifically, previous schooling, could have provided significant data.

STAAR English I Reading and STAAR English II Reading. Data indicated that the higher the TELPAS 2013 score, the higher the STAAR English II reading score. The same was true of TELPAS 2012, the higher the score, the higher the STAAR English II reading score. The differences in means were significant when considering all the participants in the study together, as well as when considering the treatment and comparison groups independently. The data are meaningful as they show that the more English a student had acquired, the better his/her performance on TELPAS 2013 reading, thus the better his/her performance on STAAR English II reading. Students who received advanced high on TELPAS 2013 reading had acquired more academic English than their peers who received ratings of beginning, intermediate, or advanced, indicating that they had acquired more CALP. Since the STAAR English II reading exam requires a high level of CALP, it stands to reason that students who had acquired the most English would perform best on the exam.

TELPAS 2012. There was a significant relationship between TELPAS 2012 reading scores and TELPAS 2013 reading scores. Not surprisingly, TELPAS 2012 reading scores were the best indicator of TELPAS 2013 reading scores. The better

students performed on TELPAS 2012 reading, the better they performed on TELPAS 2013 reading. This was the same for the treatment and comparison groups independently and together. Results indicated that when the effect of TELPAS 2012 reading was controlled for, there was not a statistically significant difference in TELPAS 2013 reading scores between the treatment and comparison groups.

One consideration when looking at means for TELPAS reading scores is that all Limited English Proficient students must take the TELPAS assessment even if they have previously scored at the advanced high level. Over the years, teachers have shared concerns that some students stop taking the test seriously in subsequent years once they have reached the advanced high level.

Another consideration is that the TELPAS reading assessments are computer-based tests. It would seem that this would be more advantageous to the group that received blended instruction as the majority of their assignments were done on the computer. However, if students had been exposed to the TELPAS at least once a year since enrolling in Texas public schools, it diminishes the potential advantage for students who received the blended method of instruction.

Language Acquisition as Measured by STAAR English II Interpretations

Research question two posed the inquiry is there a difference in the language acquisition of ELLs who take a blended ESL English class or a blended ESL reading class in comparison to ELLs who take a traditional ESL English or ESL reading class as measured by the STAAR English II reading exam. Analysis of variance indicated that although the actual scores of the group that received the traditional instructional

method ($M=1741$) were higher than the group that received the blended instructional method ($M=1731$), the difference was not statistically significant. All of the considerations for the overall TELPAS 2013 reading scores are the same for the STAAR English II reading scores. Although the n values were different, with a total of 150 scores (76 in the treatment group and 74 in the comparison group), the implications of the small n values are the same.

Additionally, another consideration is that the 2012-2013 school year was the first administration of the STAAR English II reading exam. Although TEA had released a blueprint and a test design, there was still quite a bit of uncertainty surrounding the specifics of the measure. It was new to teachers and students. This is significant because all of the teachers in the study, treatment and comparison group alike, were working to prepare students for an exam they were not familiar with and were still discovering for themselves.

The final consideration is that the STAAR English II reading scores account for only one-half of the assessment administered to students at the end of the English II course. Students must also take and pass the STAAR English II writing exam. The writing exam requires students to answer revising and editing multiple choice questions and to write three essays. Students write one expository essay, one persuasive essay, and one field tested essay that may be either expository or persuasive. It is possible that when considering both the reading and writing STAAR English II scores, statistically significant differences would have emerged.

STAAR English II Reading and the Independent Variables

Gender. Although the actual scores yielded different mean scores when considering all males ($M=1742$) and all females ($M=1730$), the differences were not statistically significant. The same was found to be true when comparing mean scores based on gender in the group that received instruction via the blended method and the group that received instruction via the traditional method. The results were not surprising as research does not identify gender as a contributing factor in second language acquisition.

Age. In reference to actual scores, STAAR English II reading scores varied by age. However, there was not a statistically significant difference on STAAR English II reading scores based on age. This was true for all participants as well as for the treatment and comparison groups independently. Further, there was not a particular pattern in mean scores based on age, i.e. the older the student the higher or lower his/her STAAR English II reading score.

TELPAS Years in U.S. Schools. Similar to the findings with TELPAS 2013, TELPAS years in U. S. schools had an effect on STAAR English II reading scores. With the exception of students with zero and three years in U.S. schools, means on STAAR English II reading were higher depending on number of years in U.S. schools. In other words, the higher the years in U.S. schools, the higher the STAAR English II reading scores. Post hoc tests indicated that there was a significant difference between students in year zero ($M=2055$) and year three ($M=1636$) with a mean difference of 419. However, there were only three students total with TELPAS years in U.S. schools at zero

in comparison to the 29 students at three years in U.S. schools. Further, the students with zero years in school had the smallest standard deviation at 176.

It was expected that the longer students were in U.S. schools, the higher their scores on standardized tests. Ideally, more time in U.S. schools equates to more time to acquire academic English. What was not expected was to have a mean score at three years in U.S. schools that was lower than means for almost all other years in U.S. schools. Also, it had the smallest SD at 181 meaning that the scores were not as spread out across the distribution as the scores for the other years in U.S. schools.

One possibility for the low mean score at year three could be that students were transitioning from BICS to CALP. Ideally, learners are right at the end of BICS acquisition and are at the brink of CALP acquisition during year three. Therefore, it is possible that students had enough language to be successful on the TELPAS reading exam, yet not enough academic language to perform well on the STAAR reading exam as it requires higher levels of CALP. Even though the students are still acquiring academic English, they are not producing it at the level required for STAAR English II reading.

Another possibility is that language acquisition may start to plateau at year three. Some researchers (Richard, 2008) indicate that often times those acquiring a second language reach a plateau when transitioning from the intermediate to the advanced level of proficiency, which occurs most often at year three. It is possible that the low means are a result of the plateau.

When comparing the treatment group to the comparison group, actual scores for students in the treatment group were higher at one, two, and five U.S. years in

schools than comparable students in the comparison group. Conversely, the students in the comparison group had higher actual scores at zero, three, four, and six years in U.S. schools in comparison to students in the treatment group at the same years in school. None of the differences however were statistically different.

Instructional Level. Results indicated that as students' instructional levels increased, so did their mean scores on STAAR English II reading. Furthermore, the differences in means were statistically significant. This pattern held true for all participants together as well as for the treatment group and the comparison group participants when considered independently. Ideally, the higher the instructional level of the student, the more English he/she had acquired. Subsequently, the STAAR English II reading exam required students to have a firm grasp of academic English. Therefore it is reasonable to expect better performance on STAAR English II reading as instructional level increases. Data indicate that these students had acquired more English so it stands to reason that they would have performed better on the STAAR English II reading exam.

Home Language and Refugee Status. As indicated in the conclusions listed about TELPAS 2013 reading and home language, the original plan was to conduct an ANOVA. However, since there were 16 home languages represented and one "other" category for those students who speak a language that was not delineated, the results would not have yielded an accurate representation of the population. Additionally, the number of students identified as refugees represented approximately 14 percent of the sample population. This number was too small to provide meaningful results.

STAAR English I Reading. There was a significant relationship between STAAR English I reading scores and STAAR English II scores. Not surprisingly, STAAR English I scores were an indicator of STAAR English II reading scores. Analysis of covariance was run to control for the effect of STAAR English I reading scores on STAAR English II reading scores. The reason was that if students in either the treatment or comparison group started the study with significantly higher or lower STAAR English I reading scores, then it would not impact the results for STAAR English II reading. Results indicated that when the effect of STAAR English I reading scores were controlled for, there was not a statistically significant difference in STAAR English II reading scores between the treatment and comparison groups.

One consideration is the impact of multiple exposure to the exam. If students who do not qualify for provisions fail the STAAR English I reading exam, they must retake the exam until they pass. The second administration takes place in the summer and the third administration takes place in the fall of the subsequent school year. Although this study only used the results from the first administration, if students took part in the second and/or third administration of the exam, it could have increased their comfort level with the organization and structure of the test.

Limitations

Although the research was carefully planned, there were some notable limitations. First, there was no true randomization used in selecting the treatment and comparison groups. As a result, it was more challenging to control for extraneous variables. Additionally, there was only one teacher in the treatment group and six in the comparison

group. It would have been more ideal to have an equal number of teachers in the treatment and comparison groups to again help control for extraneous variables. Furthermore, since a sample of convenience was used as opposed to a random sample, the generalizability of the study cannot be applied to a larger population outside of the district, it may only be suggested.

Moreover, the treatment group and the comparison group could not be matched on more than two variables because the available population was relatively small making the sample size relatively small. Therefore, the findings in the study may only be used to provide insight into contexts that are comparable to those of the study. Also, small sample sizes may attribute to inconclusive results.

Additionally, the treatment group teacher resigned at the end of April which was after both the TELPAS reading test and the STAAR English II reading test. She went to the business sector as she believed that education was not ready for the type of technological practices she wanted to implement. Her vision was far greater than the parameters in which she had to work.

Implications for Practice

The findings of the study indicate that first-line instruction is what matters most to students' learning. Regardless of instructional method, be it blended or traditional, when teachers implement salient pedagogical practices, students acquire academic English more readily. Specifically, ELLs benefit when teachers incorporate content and language objectives, provide vocabulary development, implement scaffolding, and promote academic interaction in the classroom.

While not every school and every teacher may offer an entire class using the blended model, it is important to integrate technology into the classroom in meaningful ways. Many students are tech savvy when it comes to Twitter, Facebook, and Instagram, yet are not as experienced with the uses of technology in the educational realm. Therefore, teachers must find ways to incorporate technology into the classroom in meaningful ways.

Recommendations for Future Research

One recommendation would be to repeat the study, and add student and teacher interviews about their perceptions of the effectiveness of taking a class via a blended method of instruction. Often times students' progress is not immediately seen on standardized measures, yet it does not mean that there was not progress. Furthermore, if statistically significant results are not observed, it does not mean that the class was not perceived as effective by the teacher and/or students. Therefore repeating the study and adding the qualitative piece may make a difference in findings.

Another recommendation is that the study be conducted again, but with a larger sample and with the use of true randomization. A larger sample size would allow for equal numbers of teachers in the treatment and comparison groups. True randomization could be used to select the treatment and comparison group participants as well as the teachers for each of the groups involved in the study. It would allow for the random placement of participants in either the treatment or comparison group. Further, the findings would be more generalizable with a larger sample and true randomization.

Summary

As districts and schools contend with the demands of federal and state accountability measures, there is greater onus on meeting the academic needs of all students including those in special populations, and in particular, ELLs. English-language learners have presented a challenge for districts and schools attempting to meet accountability standards. The challenge is finding ways to help ELLs acquire academic English both effectively and efficiently so that they are capable of performing satisfactorily on standardized tests. Research indicates that academic language acquisition is fundamental if ELLs are to perform well. Therefore, districts, schools, and teachers must find ways to help ELLs acquire the academic English they need to be successful both in and out of school.

References

- Agbatogun, A. (2012). Exploring the efficacy of student response system in a Sub-Saharan African country: A sociocultural perspective. *Journal of Information Technology Education Research, 11*, 249-262.
- Ariza, E. (2009). *Not for ESOL teachers: What every classroom teacher needs to know about the linguistically, culturally, and ethnically diverse student* (2nd Ed.). Boston, MA: Allyn & Bacon.
- August, D., & Shanahan, T. (2006). *Developing literacy in second-language learners: Report of the national reading panel on language-minority children and youth*. New Jersey: Lawrence Erlbaum Associates, Inc.
- Barbour, M., Brow, R., Hasler-Waters, L., Hoey, R., Hunt, J.L., Kennedy, K., Ounsworth, C., Powell, A., & Trimm, T. (2011). *Online and blended learning: A survey of policy and practice of k-12 schools around the world*. Vienna, VA: International Association for K-12 Online Learning.
- Barcroft, J. (2012). *Input-based incremental vocabulary instruction*. Alexandria, VA; TESOL International Association.
- Barr, S., Eslami, Z.R., & Joshi, R.M. (2012). Core strategies to support English language learners. *Educational Forum, 76*(1), 105-117.
- Bauer, E., & Arazi, J. (2011). Promoting literacy development for beginning English learners. *Reading Teacher, 64*(5), 383-386. doi:10.1598/RT.64.5.14.
- Bennett, S. (2007). *That workshop book: New system and structures for classrooms that reading, write, and think*. Portsmouth, NH: Heinemann.

- Center for Applied Linguistics (2013). Retrieved on June 30, 2013, from <http://www.cal.org/siop/>.
- Chung, S.F. (2012). Research-based vocabulary instruction for English language learners. *Reading Matrix: An International Online Journal*, 12(2), 105-120.
- Collier, V. (1987). Age and rate of acquisition of second language for academic purposes. *TESOL Quarterly*, 21(4). 617-641.
- Collier, V. (1989). How long? A synthesis of research of students examining long-term language minority student data on academic achievement. *TESOL Quarterly*, 23. 509-531.
- Collier, V. (1995). *Acquiring a second language for school*. Washington, DC: National Clearinghouse for Bilingual Education.
- Cummins, J. (1979a). Cognitive/academic language proficiency, linguistic interdependence the optimum age question and some other matters. *Working Papers on Bilingualism*, 19. 121-129.
- Cummins, J. (1981). The role of primary language development in promoting educational English for language minority, *Schooling and language minority students: A theoretical framework* (2ed., pp.3-48). Los Angeles: California State University, National Evaluation Dissemination and Assessment Center.
- Cummins, J. (1995). Power and pedagogy in the education of language minority students. In J. Frederickson (Ed.) *Reclaiming our voices*. Ontario, CA: California Association for Bilingual Education.
- Cummins, J. (2000). BICS and CALP. In, *Routledge Encyclopedia of Language*

- Teaching and Tips* (pp. 76-79). Taylor and Francis, Ltd. Books
- Dale, E. (1969). *Audiovisual methods in teaching*. (3rd ed.). New York, NY: Holt, Rinehart, & Winston, Inc.
- Dziuban, C., Hartman, J. & Mosaki, P. (2004). Blended learning. August 12, 2013, from <http://net.educause.edu, net.educause.edu/ir/library/pdf/erb0407.pdf>.
- Echevarria, J., Richards-Tutor, C., Chinn, V.P., & Ratleff, P. (2011). Did they get it? The role of fidelity in teaching English learners. *Journal of Adolescent and Adult Literacy*, 54(6), 425-434. doi:10.1598/JAAL.54.6.4
- Echevarria, J., Vogt, M.E., & Short, D. (2008). *Making Content Comprehensible: The SIOP Model* (3rd ed). Boston, Allyn & Bacon.
- Edelsky, C. (1996). *With literacy and justice for all: Rethinking social language and education* (2nd ed.). London: Taylor and Francis.
- Edelsky, C., Altwerger, B., & Flores, B. (1991). *Whole language: What's the difference?* Portsmouth, NH: Heinemann.
- Field, A. (2009). *Discovering statistics using SPSS* (3rd ed). Los Angeles, CA: Sage.
- Flores, S.M., Batalova, J., & Fix, M. (2012). *The educational trajectories of English language learners in Texas*. Retrieved August 1, 2013, from www.migrationpolicy.org/pubs/texasells.pdf.
- Fraenkel, J. & Wallen, N. (2008). *How design and conduct research in education* (4th ed.). New York: McGraw-Hill.
- Francis, D., Rivera, M., Lesaux, N., Kieffer, M., & Rivera, H. (2006). *Research-based*

- recommendations for serving adolescent newcomers.* Portsmouth, NH: RMC Research Corporation, Center on Instruction.
- Genesee, F. (1999). Program alternatives for linguistically diverse students: Educational practice report 1. Santa Cruz, CA; Center for Research on Education, Diversity and Excellence.
- Genesee, F., Lindholm-Leary, K., Saunders, W., & Christian, D. (2006). *Educating language learners: A synthesis of research evidence.* New York: Cambridge University Press.
- Hakuta, K., Butler, Y.G., & Witt, D. (2000). *How long does it take English learners to attain proficiency?* UC Berkeley: University of California Linguistic Minority Research Institute.
- Horn, M.B., Staker, H., & Innosight, I. (2011), The rise of K-12 blended learning. *Innosight Institute.*
- Interaction (n.d.). In *Merriam-Webster Dictionary Online*. Retrieved from <http://www.merriam-webster.com>
- Krashen, S. (1981). *Second language acquisition and second language learning.* Oxford: Pentagon.
- Krashen, S. (1985). *The input hypothesis: Issues and implications.* New York, NY: Longman Publishing.
- Krashen, S.D., Long, M. A., & Scarcella, R.C. (1979). Age, rate, and eventual attainment in second language acquisition. *TESOL Quarterly*, 13(4). 573-582.
- Lim, D.H., Morris, M.L., & Kupritz, V.W. (2007). Online vs. blended learning:

Journal of Asynchronous Learning Network, 11(2), 27-42.

- Marzano, R.J., & Pickering, D. J. (2005). *Building academic vocabulary: Teacher's manual*. Alexandria, VA: Association for Supervision and Curriculum Development.
- McLaughlin, B., Rossman, T., & McLeod, B. (1984). Second language learning: An information-processing perspective. *Language Learning*, 33(2). 309-332.
- Means, B. (2010). *Evidence-based practices in online learning*. Washington, DC: U.S. Department of Education, Office of Planning, Evaluation, and Policy Development.
- Means, B., Toyama, Y., Murphy, R., Bakia, M., & Jones, K. (2010). *Evaluation of evidence-based practices in online learning: A meta-analysis and review of online learning studies*. Washington, DC: U.S. Department of Education, Office of Planning, Evaluation, and Policy Development.
- Mitchell & Myles (2004). *Second language learning theories*. (2nd ed.). London: Hodder Arnold.
- Morrison, D.M., & Low, G. (1983). Monitoring and the second language learner. In J. Richards and Schmidt eds. *Language and communication*. London: Longmann.
- National Clearinghouse for English Language Acquisition and Language Instruction Educational Programs. (2010). *The growing numbers of limited English students*. Retrieved June 25, 2013, from <http://www.ncela.gwu.edu/policy.states/stateposter/pdf>.
- National Center of Education Statistics (2013). Biannual reports to congress on the

implementation of the Title III state formula grant program school years 2006-2008. Retrieved on August 1, 2013, from <http://www.ed.gov>.

National reading panel teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction (2000, April 13). *U.S. Department of Health and Human Services: National Institutes of Health*. Retrieved March 12, 2012, from <http://www.nichd.nih.gov/publications/pubs/nrp/pages/smallbook.aspx>

Pawan, F. (2008). Content-area teachers and scaffolded instruction for English language learners. *Teaching and Teacher Education*, 24. 1450-1462.

Picciano, A.G., & Seaman, J. (2009). K-12 online learning: A follow-up of the survey of U.S. school direct administrators. Sloan Consortium.

Region 13 Education Service Center (2013). *Professional development and appraisal system*. Retrieved on June 25, 2013, from <http://ww4.esc13.net/pdas/pdf>.

Richard, J.C. (2008). *Moving beyond the plateau: From intermediate to advanced levels in learning language*. New York, NY: Cambridge University Press.

Richard-Amato, P.A. (2003). *Making it happen: From interactive to participatory language teaching theory and practice*. (3rd ed.). White Plains, NY: Pearson Education, Inc.

Rodrigo, V., Krashen, S., & Gribbons, B. (2004). The effectiveness of two comprehensible-input approaches to foreign language instruction at the intermediate level. *System*, 32(1). 53.

Rustipa, K. (2011). *Contrastive analysis, error analysis, interlanguage and the*

- implication to language teaching*. Retrieved October 15, 2013, from http://www.polines.acid/ragam/index_files/jurnalragam/paper_3apr_2011.pdf.
- Saye, J. W., & Brush, T. (2002). Scaffolding critical reasoning about history and social issues in multimedia-supported learning environments. *Educational Technology Research and Development* 50(3), 77-96.
- Short, D., & Echevarria, J. (2004). Teacher skills to support English language learners, *Educational Leadership*, 62(4), 8-13.
- Staker, H., & Innosight, I. (2011). The rise of K-12 blended learning: Profiles of emerging models. *Innosight Institute*.
- Texas Education Agency (2007). *English language proficiency standards*. Retrieved April 3, 2013, from <http://www.tea.state.tx.us>.
- Texas Education Agency (2010). *The state of Texas assessment of academic readiness (STAAR): A new assessment model*. Retrieved June 30, 2013, from <http://www.tea.state.tx.us>. ProposedNewAssessmentModel[1].pdf.
- Texas Education Agency (2011a). *2011 Accountability manual*. Retrieved on June 25, 2013, from <http://ritter.tea.state.tx.us/perfreport/account/2011/manual/>.
- Texas Education Agency (2011b). *Educator guide to TELPAS grades k-12*. Retrieved on June 30, 2013, from <http://www.tea.state.tx.us>. TELPAS-InfoGuide2012.pdf.
- Texas Education Agency (2011c). *Title I school improvement program: Part I LEA and school improvement 2011-2012*. Retrieved on July 9, 2013, from www.tea.state.tx.us 2011-2012_Guidance_Part1_School_LEA-Improv.pdf
- Texas Education Agency (2012a). *2012-2013 STAAR decision-making guide for LPACs*.

- Retrieved May 15, 2013, from <http://tea.state.tx.us>. LPAC-STAAR-Decisions-MakingGuide2013.pdf.
- Texas Education Agency (2012b). *Fall ELL assessment update*. Retrieved May 15, 2013, from <http://www.tea.state.tx.us>. ELL-AssessmentUpdate-TETN15880-Fall2012.pptx.
- Texas Education Agency (2012c). *LPAC framework manual 2012-2013*. Retrieved on June 23, 2013, from <http://porta.esc20.net>.
- Texas Education Agency (2012d). *Technical Digest*. Retrieved June 24, 2013, from <http://www.tea.state.tx.us>. digest12-ch7.pdf
- Texas Education Agency (2013a). *2013 Accountability manual*. Retrieved on August 1, 2013, from <http://ritter.tea.state.tx.us/perfreport/account/2013/manual/>.
- Texas Education Agency (2013b). *Accountability system for 2013 and beyond technical description: Performance index indicators and index construction*. Retrieved on August 1, 2013, from <http://ritter.tea.state.tx.us/perfreport/account/2013/>.
- Texas Education Agency (2013c). *English language learners web portal*. Retrieved May 15, 2013, from <http://elltx.org>.
- Texas Education Agency (2013d). *State of Texas assessments of academic readiness assessments: Standard setting technical report*. Retrieved on July 15, 2013, from www.tea.state.tx.us STAAR-TechRpt-Combined.pdf
- Thomas, W.P., & Collier, V. (1997). *School effectiveness for language minority students*. Washington, DC: National Clearinghouse for Bilingual Education.
- Thomas, W.P., & Collier, V., (2002). A national study of school effectiveness for

language minority students' long-term academic achievement.

Torgensen, J., Houston, D., Rissman, L., Decker, S., Roberts, G., Vaughn, S., Wexler, J.,

Francis, D., Rivera, M., & Lesaux, N. (2007). *Academic literacy instruction for*

adolescents. Portsmouth, NH: RMC Research Corporation, Center on

Instruction.

Vygotsky, L. (1978). *Mind in society*. Cambridge, MA: MIT.

Wallace, C. (2008). Vocabulary: The key to teaching English language learners to read.

Reading Improvement. 36-39.

Walqui, A. (2006). Scaffolding instruction for English language learners: A conceptual

framework. *International Journal of Bilingual Education and Bilingualism*, 9 (2).

159-180.

Watson, J. (2008). Blending learning: The convergence of online and face-to-face

education. Vienna, VA: North American Council for Online Learning.

Watson, J., Murin, A., Vashaw, L. Gemin, B. & Rapp, C. (2010). Keeping pace with K-

12 online learning: An annual review of policy and practice. Retrieved

September 2, 2013, from

kpk12.com/cms/wpcontent/uploads/KeepingPace2012.pdf.

Wiley, T.G. (1996). *Literacy and language diversity in the United States*. Washington,

D.C.: Center for Applied Linguistics and Delta Systems.

Wood, D., Bruner, J., & Ross, G. (1976). The role of tutoring in problem solving. *Journal*

of Child Psychology and Psychiatry, 17, 89-10

Zacher-Pandya, J. (2011). Overtested: *How high-stakes accountability fails English*

language learners. New York, NY: Teachers College Press.

Appendix A
TELPAS Ratings

TELPAS Reading Levels	Degree of Linguistic Accommodation Applied to Passage and Item Development	
Advanced High	Minimal	Minimal linguistic accommodation; texts highly comparable to those written for native English speakers.
Advance	Moderate	Occasional picture support; contextual aids and organizational features support comprehension of longer texts on both familiar and unfamiliar language arts and content area topics.
Intermediate	Substantial	Frequent picture support short texts written primarily on familiar topics; commonly used, everyday English and routine academic English.
Beginning	Extensive	Maximum picture support; short texts that require comprehension of words, phrases, and short sentences that use the type of high-frequency, concrete vocabulary first acquired by learners of a second language.

Educator Guide to TELPAS, p. 36.

Appendix B

Traditional Instructional Method Lesson Plans

Big Idea/Lesson Understanding: To Discover the impact of cultural traditions throughout the Generations	Guiding Questions: Why is it necessary to pass cultural traditions from one generation to the other? How does the immigrant experience change over generations? Why are traditions important/
Content Objective/TEKS: Students will: Analyze guiding questions to prepare for the theme: "traditions" and immigration. Analyze how elements of a short story interact creating believable characters. Read: The Keeping Quilt" by Patricia Pallicio.	Language Objective: Students will: Think about Traditions that relate to their culture. Discuss and share connections Brainstorm ideas using a Thinking Map Frame the map by writing an answer in English: Traditions come from our..... and
Anticipatory Set: (Opener, Engage, Hook) Students will: View visual images on the P. board from different cultures. Discuss and comment on why these are important in our culture. Write about something they have at home that is an object illustrating part of their culture.	Checking for Understanding: (Conferring, Questioning, Feedback, occurs throughout lesson) Teacher will monitor and check/answer questions asked by students.
Instructional Input & Modeling: (Mini-Lesson, Explore, Explain, Line, "I-Do") Teacher will explain or review Thinking maps. We will discuss why we use the circle map.	
Guided Practice: (Catches, Elaborate, Line, "We-Do") Teacher will demonstrate examples of what to write on circle map by using the United States as an example	
Independent Practice: (Work-Time, Evaluate, Sink, "You-Do") Students will: Complete their Thinking maps by writing about where traditions come from in their culture. Answer the question on the frame.	Student Use of Technology: Use of the P. Board.
Closure: Show students a Picture of a quilt and ask them to answer the question: Why might this be an important tradition that is passed from one generation to the other.	

Big Idea/Lesson Understanding: Influences through the media	Guiding Questions: How does the media shape the way teens think?
Content Objective/TEKS: Increase comprehension/fluency Understanding of persuasive techniques	Language Objective: Students will jigsaw read selected text, write main points, speak and listen to table partners and respond orally in classroom discussion
Anticipatory Set: (Opener, Engage, Hook) Students will look at “visual” displayed on promethean board- Cigarette package cover of “Kool” cigarettes (rapper). Students will write response to visual on whiteboard and display to table partner	Checking for Understanding: (Conferring, Questioning, Feedback, occurs throughout lesson) Monitor students throughout the lesson while jigsaw reading Answer any questions giving ample “wait-time” for answers and/or thoughtful questions Give feedback as needed
Instructional Input & Modeling: (Mini-Lesson, Explore, Explain, Line, “I-Do”) TTW: Explain the use of persuasive techniques in the media. Will show an example with the use of promethean board – display five multiple choice questions pertaining to teens and cigarette smoking. Teacher will read and explain several questions and discuss the persuasive techniques used to target teenagers. This includes influential wording and visuals. Teacher will show several visuals of hip-hop characters and the various cigarette package covers marketed to target teens.	
Guided Practice: (Catches, Elaborate, Line, “We-Do”) Students are given short text “How do Tobacco Companies market to teenagers?” Students are assigned in groups of three or four. Each group will be assigned a “chunk” of the selection to read. In each group there is a recorder, speaker, time-keeper. Students are given 25 minutes to read their portion of the selection and write down the main points including persuasive techniques used. When finished the recorder will write down the main points of the selection on chart paper. The chart paper will be posted around the room and used for a gallery walk. The speaker will explain his/her group’s main ideas during the walk.	
Independent Practice: (Work-Time, Evaluate, Sink, “You-Do”) During the gallery-walk each student will write down one or two main ideas/persuasive techniques he/she learned (off of the chart paper).	Student Use of Technology: Promethean board
Closure: Exit ticket: write one persuasive technique used to market to teens from the gallery walk.	

Big Idea/Lesson Understanding: To Discover the impact of cultural traditions throughout the Generations	Guiding Questions: Why is it necessary to pass cultural traditions from one generation to the other? How does the immigrant experience change over generations? Why are traditions important/?
Content Objective/TEKS: Students will: Analyze guiding questions to prepare for the theme: "traditions" and immigration. Analyze how elements of a short story interact creating believable characters. Read: The Keeping Quilt" by Patricia Pallicio.	Language Objective: Students will: Review and retell story details Define new voc. Read and comprehend part 1 of the story. Write answers to comprehension questions
Anticipatory Set: (Opener, Engage, Hook) Look at pictures of immigrants: Why do people come to the USA? Discuss.	Checking for Understanding: (Conferring, Questioning, Feedback, occurs throughout lesson) Teacher will monitor and check/answer questions asked by students.
Instructional Input & Modeling: (Mini-Lesson, Explore, Explain, Line, "I-Do") Teacher will go over and review voc from part 1. Discuss the question: How do we know that the family came to New York for work.	
Guided Practice: (Catches, Elaborate, Line, "We-Do") Teacher will explain how the students will read in groups and use the summary strategy after each paragraph .	
Independent Practice: (Work-Time, Evaluate, Sink, "You-Do") Students will: Read part one silently Reread Part one in groups Use the say something or summary strategy to talk together about each paragraph. Write answers to comprehension questions on scaffolded lesson.	Student Use of Technology: Use of P. Board
Closure: Exit Ticket: Complete the sentence: the Quilt reminds the family of their homeland because.....	

Figurative Language

Student Friendly Objectives:

TLW...

1. explain figurative language;
2. distinguish between figurative and non-figurative writings;
3. identify and differentiate among the different figures of speech; and
4. demonstrate an eye and ear for figurative language as they listen keenly to songs and explain works in songs.

Materials:

1. Teacher-compiled songs
2. Framed quotations
3. Supplementary material: WRITE AHEAD book

Setting the Stage:

1. Students decide whether the framed quotes they have picked are figurative or not.
2. They explain their choices.

Student Actions/Engagement:

1. They listen to old and new songs intently. They pay attention to their lyrics.
2. They identify metaphorical language, allusions, and other rhetorical devices of which these songs are replete.

Teacher Actions/Instructional Strategies:

1. Teacher groups the figures of speech to facilitate mastery of them.
2. Teacher adopts the SHAAAMPOOO mnemonic device.

Closure/Assessment:

A figurative classroom tour takes place as students fill their SHAAAMPOOOs with examples from around the classroom.

Homework:

Relate figurative language to poetry.

English II Lesson Plan					
Content Objectives	Resources	Opening Structure	Mini Lesson	Work time	Debrief
Learn and create examples of expository strategies	Expository writing graphic organizer	Introduce elements of exposition.	Teacher reviews expository strategies and examples.	Students discuss expository writing strategies, and write a student example of each using the subject "good health".	View student examples under document viewer.

Subjunctive mood, semicolons, review of persuasive and literary essays

Student Friendly Objectives:

TLW...

1. use correct verb forms in hypothetical statements;
2. use semicolons correctly;
3. draft persuasive and literary essays;
4. identify strengths and weaknesses based on feedback from graded papers given back; and
5. answer reviewers and defend their answers.

Materials:

1. Eric Clapton's song If I Were a King
2. Handouts
3. Graded persuasive and literary essays
4. Online practice on semicolons

Homework:

Practice persuasive writing by defending your answer on: Who is to blame for the death of Romeo and Juliet?

Appendix C

Blended Instructional Method Lesson Plans

English II						
Day	Objectives	Resources	Opening Structure	Mini Lesson	Work time	Debrief
Mon.	Identify elements of a good expository essay	Effects of watching too much TV	Identify the three levels of text structure- text level, sentence level, word level	Breakdown thesis. I.D. where the info found in essay	Look at word, sentence, and text levels. Explain why and how used	Students share their findings. Annotate their copies.
Tues.	Revise the essay paying attention to word and sentence level	Texas Writing Coach (online)	Remind students of sentence and word level structures and how they improve writing	Model with sentences chosen from the students writing	Revise first paragraph of essay. Justify why changes made.	Students share their revisions and justifications.
Wed.	Revise the essay paying attention to text level structure	Texas Writing Coach (online)	Remind students of the overall text structure of the expository essay	Model text level structures with model essay	Students ensure that their essay has all the elements.	Students share changes they made.
Thurs.	Student will evaluate writing using the editing process	Texas Writing Coach (online)	Explain the editing process	Take a student's essay and annotate with the class	Annotate each other's essays. Use ebook to help them.	Share student examples.
Fri.	Student will evaluate writing using the editing process	Essay Sample Texas Writing Coach (online)	Essay Sample Revising and editing procedure	Take a student's essay and annotate with the class	Students complete the editing process of own essay.	Evaluate the revising and editing process.

Appendix D

Descriptive Statistics for All Variables

Descriptive Statistics All Participants

		Gender	Home Language	Instructional Level	Age
N	Valid	194	194	194	194
	Missing	0	0	0	0
Mean		1.5670	10.8505	2.875	15.9563
Median		2.0000	11.0000	3.0000	16.0000
Mode		2.00	11.00	3.00	16.00
Std. Deviation		.49677	3.94172	.94685	.92904
Variance		.247	15.537	.897	.863
Skewness		-2.73	-.615	-.490	.719
Kurtosis		-1.946	.114	-.648	.693

		Refugee Status	Years in US School	STAAR9	STAAR10
N	Valid	194	194	145	150
	Missing	0	0	49	44
Mean		.1340	3.3763	1633.33	1736.08
Median		.0000	3.0000	1631.00	1721.50
Mode		.00	6.00	1750.00	1623.00
Std. Deviation		.34156	1.87087	164.102	212.208
Variance		.117	3.500	26929.57	45032.30
Skewness		2.165	.030	.121	.303
Kurtosis		2.717	-1.232	-.332	.135

		TELPAS 2012	TELPAS 2013
N	Valid	172	190
	Missing	22	4
Mean		2.9593	3.2579
Median		3.0000	3.0000
Mode		3.00	4.00
Std. Deviation		.97547	.81767
Variance		.952	.669
Skewness		-.606	-.685
Kurtosis		-.633	-.654

Descriptive Statistics Treatment Group

		Gender	Home Language	Instructional Level	Age
N	Valid	91	91	91	91
	Missing	0	0	0	0
Mean		1.5714	11.0769	3.0000	15.8242
Median		2.0000	11.0000	3.0000	16.0000
Mode		2.00	11.00	3.00	16.00
Std. Deviation		.49761	4.06948	.81650	.82453
Variance		.248	16.561	.667	.680
Skewness		-2.94	-.563	-.501	.583
Kurtosis		-1.957	.115	-.216	.001

		Refugee Status	Years in US School	STAAR9	STAAR10
N	Valid	91	91	77	76
	Missing	0	0	14	15
Mean		.1538	3.6484	1611.31	1731.32
Median		.0000	4.0000	1631.00	1690.00
Mode		.00	6.00	1750.00	1623.00
Std. Deviation		.36280	1.77872	231.124	204.546
Variance		.132	3.164	53418.402	41839.046
Skewness		1.951	-.213	-3.332	.576
Kurtosis		1.847	-1.087	21.481	.396

		TELPAS 2012	TELPAS 2013
N	Valid	84	88
	Missing	7	3
Mean		3.1071	3.3636
Median		3.0000	3.0000
Mode		3.00	4.00
Std. Deviation		.83620	.68114
Variance		.699	.464
Skewness		-.713	-.505
Kurtosis		.000	-.698

Descriptive Statistics Comparison Group

		Gender	Home Language	Instructional Level	Age
N	Valid	103	103	103	103
	Missing	0	0	0	0
Mean		1.5631	10.6505	2.7670	16.0680
Median		2.0000	11.0000	3.0000	16.0000
Mode		2.00	11.00	3.00	16.00
Std. Deviation		.49843	3.83408	1.04040	1.00257
Variance		.248	14.700	1.082	1.005
Skewness		-2.58	-.705	-.369	.695
Kurtosis		-1.972	.412	-1.022	.709

		Refugee Status	Years in US School	STAAR9	STAAR10
N	Valid	103	103	68	74
	Missing	0	0	35	29
Mean		.1165	3.1359	1638.94	1740.97
Median		.0000	3.0000	1641.50	1775.00
Mode		.00	2.00	1490.00 ^b	1551.00
Std. Deviation		.32240	1.92531	172.665	221.092
Variance		.104	3.707	29813.250	48882.027
Skewness		2.426	.257	-.061	.077
Kurtosis		3.963	-1.203	-.524	.021

		TELPAS 2012	TELPAS 2013
N	Valid	88	102
	Missing	15	1
Mean		2.8182	3.1667
Median		3.0000	3.0000
Mode		4.00	4.00
Std. Deviation		1.07794	.91287
Variance		1.162	.833
Skewness		-.416	-.578
Kurtosis		-1.103	-1.004

Appendix E
Home Language Frequencies

All Participants

Language	Frequency	Percent	Valid Percent	Cumulative Percent
Arabic	3	1.5	1.5	1.5
Bengali	1	.5	.5	2.1
Burmese	17	8.8	8.8	10.8
Cantonese	3	1.5	1.5	12.4
Farsi	1	.5	.5	12.9
French	2	1.0	1.0	13.9
Hakka	2	1.0	1.0	14.9
Mandarin	11	5.7	5.7	20.6
Nepali	4	2.1	2.1	22.7
Somali	2	1.0	1.0	23.7
Spanish	96	49.5	49.5	73.2
Swahili	5	2.6	2.6	75.8
Taiwanese	2	1.0	1.0	76.8
Urdu	2	1.0	1.0	77.8
Yorubu	1	.5	.5	78.4
Vietnamese	37	19.1	19.1	97.4
Other	5	2.6	2.6	100.0
Total	194	100.0	100.0	

Treatment Group

Language	Frequency	Percent in Type of Group	Percent within Home Language	Cumulative Percent
Arabic	1	1.1	33.3	0.5
Bengali	1	1.1	100	0.5
Burmese	6	6.6	35.3	3.1
Cantonese	3	3.3	100	1.5
Farsi	1	1.1	100	0.5
French	0	0	0	0
Hakka	2	2.2	100	1.0
Mandarin	5	5.5	45.5	2.6
Nepali	3	3.3	75	1.5
Somali	2	2.2	100	0
Spanish	37	40.7	38.5	19.1
Swahili	4	4.4	80.0	2.1
Taiwanese	1	1.1	50.0	0.5
Urdu	1	1.0	50.0	0.5
Yorubu	1	1.1	100	0.5
Vietnamese	19	20.9	51.4	9.8
Other	4	4.4	80.0	2.1
Total	91	100	46.9	46.9

Comparison Group

Language	Frequency	Percent in Type of Group	Percent within Home Language	Cumulative Percent
Arabic	2	1.9	66.7	1.0
Bengali	0	0	0	0
Burmese	11	10.7	64.7	5.7
Cantonese	0	0	0	0
Farsi	0	0	0	0
French	2	1.9	100	1.0
Hakka	0	0	0	0
Mandarin	6	5.8	54.5	3.1
Nepali	1	1.0	25.0	0.5
Somali	0	0	0	0
Spanish	59	57.3	61.5	30.4
Swahili	1	1.0	20.0	0.5
Taiwanese	1	1.0	50.0	0.5
Urdu	1	1.0	50.0	0.5
Yorubu	0	0	0	0
Vietnamese	18	17.5	48.6	9.3
Other	1	1.0	20.0	0.5
Total	103	100	53.1	53.1

Appendix F

Tests of Between Subjects Dependent Variable TELPAS 2013

Tests of Between-Subjects Effects Gender

Dependent Variable: TELPAS 2013

Source	Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power ^b
Corrected Model	1.173 ^a	1	1.173	1.767	.186	.009	1.762	.252
Gender	1.173	1	1.173	1.762	.186	.009	1.762	.252
Error	125.190	188	.666					
Total	2143.000	190						
Corrected Total	126.363	189						

a. R Squared = .009 (Adjusted R Squared = .004); b. Computed using alpha = .05

Tests of Between-Subjects Effects Instructional Level

Dependent Variable: TELPAS 2013

Source	Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power ^b
Corrected Model	52.156	3	17.385	43.576	.000	.413	130.728	1.000
Instructional Level	52.156	3	17.385	43.576	.000	.413	130.728	1.000
Error	74.207	186	.399					
Total	2143.000	190						
Corrected Total	126.363	189						

a. R Squared = .413 (Adjusted R Squared = .403); b. Computed using alpha = .05

Tests of Between-Subjects Effects Age

Dependent Variable: TELPAS 2013

Source	Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power ^b
Corrected Model	10.467 ^a	5	2.093	3.324	.007	.083	16.618	.893
Age	10.467	5	2.093	3.324	.007	.083	16.618	.893
Error	115.896	184	.630					
Total	2143.000	190						
Corrected Total	126.363	189						

a. R Squared = .083 (Adjusted R Squared = .058); b. Computed using alpha = .05

Tests of Between-Subjects Effects Years in U.S. Schools

Dependent Variable: TELPAS 2013

Source	Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power ^b
Corrected Model	4.688 ^a	6	.781	1.175	.321	.037	7.051	.456
YrsinSchool	4.688	6	.781	1.175	.321	.037	7.051	.456
Error	121.675	183	.665					
Total	2143.000	190						
Corrected Total	126.363	189						

a. R Squared = .037 (Adjusted R Squared = .006); b. Computed using alpha = .05

Tests of Between-Subjects Effects TELPAS 2012

Dependent Variable: TELPAS 2013

Source	Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power ^b
Corrected Model	60.765 ^a	3	20.255	64.957	.000	.543	194.870	1.000
TELPAS2012	60.765	3	20.255	64.957	.000	.543	194.870	1.000
Error	51.139	164	.312					
Total	1822.000	168						
Corrected Total	111.905	167						

a. R Squared = .543 (Adjusted R Squared = .535); b. Computed using alpha = .05

Tests of Between-Subjects Effects STAAR9

Dependent Variable: TELPAS 2013

Source	Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power ^b
Corrected Model	39.716 ^a	35	1.135	2.647	.000	.466	92.635	1.000
STAAR9	39.716	35	1.135	2.647	.000	.466	92.635	1.000
Error	45.446	106	.429					
Total	1621.00	142						
Corrected Total	85.162	141						

a. R Squared = .466 (Adjusted R Squared = .290); b. Computed using alpha = .05

Tests of Between-Subjects Effects STAAR10

Dependent Variable: TELPAS 2013

Source	Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power ^b
Corrected Model	47.567 ^a	35	1.359	3.873	.000	.545	135.548	1.000
STAAR10	47.567	35	1.359	3.873	.000	.545	135.548	1.000
Error	39.654	113	.351					
Total	1745.000	149						
Corrected Total	87.221	148						

a. R Squared = .545 (Adjusted R Squared = .405); b. Computed using alpha = .05

Appendix G

ANOVA Descriptive Statistics by Group TELPAS 2013

Descriptive Statistics Gender
Dependent Variable: TELPAS 2013

Type of Group	Gender	Mean	Std. Deviation	N
Treatment	Females	3.237	.6339	38
	Males	3.460	.7060	50
Comparison	Females	3.111	.9101	45
	Males	3.210	.9207	57
Total	Females	3.169	.7936	83
	Males	3.327	.8330	107

Descriptive Statistics Instructional Level
Dependent Variable: TELPAS 2013

Type of Group	Instructional Level	Mean	Std. Deviation	N
Treatment	Beginning	2.250	.5000	4
	Intermediate	3.000	.7670	18
	Advanced	3.415	.4988	41
	Advanced High	3.720	.6138	25
Comparison	Beginning	1.933	.5936	15
	Intermediate	2.727	.8270	22
	Advanced	3.457	.7005	35
	Advanced High	3.767	.5040	30
Total	Beginning	2.000	.5774	19
	Intermediate	2.850	.8022	40
	Advanced	3.434	.5963	76
	Advanced High	3.746	.5517	55

Descriptive Statistics Age

Dependent Variable: TELPAS 2013

Type of Group	Age	Mean	Std. Deviation	N
Treatment	14	4.000		1
	15	3.531	.6214	32
	16	3.205	.6951	39
	17	3.308	.7511	13
	18	3.667	.5774	3
Comparison	14	3.000	1.414	.2
	15	3.483	.7378	29
	16	3.317	.8786	41
	17	2.782	.9980	23
	18	2.250	.5000	4
Total	19	2.333	.5774	3
	14	3.333	1.155	3
	15	3.509	.6739	61
	16	3.263	.7915	80
	17	2.972	.9407	36
	18	2.857	.8997	7
	19	2.333	.5774	3

Descriptive Statistics Years in U.S. Schools

Dependent Variable: TELPAS 2013

Type of Group	Years in U.S. Schools	Mean	Std. Deviation	N
Treatment	.00	4.000	.0000	2
	1.00	3.539	.6602	13
	2.00	3.000	.7071	9
	3.00	3.118	.7812	17
	4.00	3.313	.6021	16
	5.00	3.571	.5136	14
	6.00	3.364	.7175	17
Comparison	.00	3.333	1.0328	6
	1.00	2.800	.8619	15
	2.00	3.154	.9572	26
	3.00	3.059	.8994	17
	4.00	3.286	.9512	7
	5.00	3.100	.9944	10
	6.00	3.476	.8136	21
Total	.00	3.500	.9258	8
	1.00	3.114	.8483	28
	2.00	3.088	.9000	35
	3.00	3.304	.8300	34
	4.00	3.378	.7029	23
	5.00	3.375	.7697	24
	6.00	3.474	.7618	38

Appendix H

Tests of Between Subjects STAAR10

Tests of Between-Subjects Effects Gender

Dependent Variable: STAAR 10

Source	Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power ^b
Corrected Model	5423.195 ^a	1	5423.195	.120	.730	.001	.120	.064
Gender	5423.195	1	5423.195	.120	.730	.001	.120	.064
Error	6704389.845	148	45299.931					
Total	458805878.0	150						
Corrected Total	6709813.040	149						

a. R Squared = .001 (Adjusted R Squared = -.006); b. Computed using alpha = .05

Tests of Between-Subjects Effects Instructional Level

Dependent Variable: STAAR 10

Source	Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power ^b
Corrected Model	1937464.67 ^a	3	645821.555	19.758	.000	.289	59.273	1.000
Instructional Level	1937464.666	3	645821.555	19.756	.000	.289	59.273	1.000
Error	4772348.374	146	32687.318					
Total	458805878.0	150						
Corrected Total	6709813.040	149						

a. R Squared = .289 (Adjusted R Squared = .274); b. Computed using alpha = .05

Tests of Between-Subjects Effects Age

Dependent Variable: STAAR 10

Source	Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power ^b
Corrected Model	167518.458 ^a	4	41879.614	.928	.449	.025	3.713	.289
Age	167518.458	4	41879.614	.928	.449	.025	3.713	.289
Error	6542294.562	145	45119.273					
Total	458805878.0	150						
Corrected Total	6709813.040	149						

a. R Squared = .025 (Adjusted R Squared = -.002); b. Computed using alpha = .05

Tests of Between-Subjects Effects Years in U.S. Schools

Dependent Variable: STAAR 10

Source	Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power ^b
Corrected Model	659559.884 ^a	6	1099926.647	2.598	.020	.098	15.589	.841
YrsinSchool	659559.884	6	109926.647	2.598	.020	.098	15.589	.841
Error	6050253.156	143	42309.463					
Total	458805878.0	150						
Corrected Total	6709813.040	149						

a. R Squared = .098 (Adjusted R Squared = .060); b. Computed using alpha = .05

Tests of Between-Subjects Effects TELPAS 2012

Dependent Variable: STAAR 10

Source	Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power ^b
Corrected Model	2817108.09 ^a	3	939036.031	40.411	.000	.473	121.232	1.000
TELPAS2012	2817108.094	3	939036.031	40.411	.000	.473	121.232	1.000
Error	3137042.424	135	23237.351					
Total	419010762.0	139						
Corrected Total	5954150.518	138						

a. R Squared = .473 (Adjusted R Squared = .461); b. Computed using alpha = .05

Tests of Between-Subjects Effects TELPAS 2013

Dependent Variable: STAAR 10

Source	Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power ^b
Corrected Model	2771368.11 ^a	3	923789.372	36.453	.000	.430	109.358	1.000
TELPAS2013	2771368.115	3	923789.372	36.453	.000	.430	109.358	1.000
Error	3674624.033	145	25342.235					
Total	453752374.0	149						
Corrected Total	6445992.148	148						

a. R Squared = .098 (Adjusted R Squared = .060); b. Computed using alpha = .05

Tests of Between-Subjects Effects STAAR9
 Dependent Variable: STAAR 10

Source	Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power ^b
Corrected Model	3396266.71 ^a	33	102917.173	4.053	.000	.565	133.757	1.000
STAAR9	3396266.706	33	102917.173	4.053	.000	.565	133.757	1.000
Error	2615306.375	103	25391.324					
Total	414687129.0	137						
Corrected Total	6011573.080	136						

a. R Squared = .565 (Adjusted R Squared = .426); b. Computed using alpha = .05

Appendix I

ANOVA Descriptive Statistics by Group STAAR 10

Descriptive Statistics Gender
Dependent Variable: STAAR 10

Type of Group	Gender	Mean	Std. Deviation	N
Treatment	Females	1717.486	226.137	35
	Males	1743.122	186.201	41
Comparison	Females	1742.000	199.160	34
	Males	1740.100	240.693	40
Total	Females	1729.565	212.0624	69
	Males	1741.630	213.495	81

Descriptive Statistics Instructional Level
Dependent Variable: STAAR 10

Type of Group	Instructional Level	Mean	Std. Deviation	N
Treatment	Beginning	1536.000	66.803	4
	Intermediate	1609.231	201.123	13
	Advanced	1700.629	158.731	35
	Advanced High	1874.750	198.846	24
Comparison	Beginning	1445.000	143.882	6
	Intermediate	1608.583	226.190	12
	Advanced	1743.074	159.176	27
	Advanced High	1855.034	200.890	29
Total	Beginning	1481.400	123.276	10
	Intermediate	1608.920	208.984	25
	Advanced	1719.113	159.038	62
	Advanced High	1863.962	198.287	53

Descriptive Statistics Age
Dependent Variable: STAAR 10

Type of Group	Age	Mean	Std. Deviation	N
Treatment	14	2000.000		1
	15	1758.241	196.069	29
	16	1707.824	211.530	34
	17	1710.319	206.205	12
Comparison	14	1817.000		1
	15	1700.720	226.051	25
	16	1804.910	231.592	33
	17	1666.786	163.515	14
	18	1600.000		1
Total	14	1908.500	129.401	2
	15	1731.611	210.445	54
	16	1755.642	225.320	67
	17	1686.923	181.946	26
	18	1600.00		1

Descriptive Statistics Years in U.S. School
Dependent Variable: STAAR 10

Type of Group	Years in U.S. Schools	Mean	Std. Deviation	N
Treatment	.00	1959.000	80.610	2
	1.00	1766.889	219.772	9
	2.00	1766.222	300.320	9
	3.00	1592.143	136.537	14
	4.00	1733.846	249.316	13
	5.00	1791.846	167.223	13
	6.00	1733.750	130.299	16
Comparison	.00	2248.000		1
	1.00	1632.333	230.240	6
	2.00	1765.353	234.326	17
	3.00	1677.200	210.268	15
	4.00	1763.857	166.118	7
	5.00	1711.111	309.986	9
	6.00	1782.842	154.114	19
Total	.00	2055.333	176.322	3
	1.00	1713.067	226.247	15
	2.00	1765.654	252.989	26
	3.00	1636.138	180.645	29
	4.00	1744.300	219.517	20
	5.00	1758.818	232.886	22
	6.00	1760.400	143.805	35

Descriptive Statistics TELPAS 2012 Score Ratings
 Dependent Variable: STAAR 10

Type of Group	TELPAS 2012 Rating	Mean	Std. Deviation	N
Treatment	Beginning	1536.000	66.803	4
	Intermediate	1540.857	161.601	8
	Advanced	1675.394	136.599	33
	Advanced High	1889.539	197.752	26
Comparison	Beginning	1445.000	143.882	6
	Intermediate	1522.152	125.426	11
	Advanced	1681.217	138.008	23
	Advanced High	1880.179	156.717	28
Total	Beginning	1491.400	123.276	10
	Intermediate	1530.053	137.788	19
	Advanced	1677.786	135.953	56
	Advanced High	1884.685	176.012	54

Descriptive Statistics TELPAS 2013 Score Ratings
 Dependent Variable: STAAR 10

Type of Group	TELPAS 2013 Rating	Mean	Std. Deviation	N
Treatment	Beginning			
	Intermediate	1516.875	143.396	8
	Advanced	1658.387	189.783	31
	Advanced High	1827.417	196.847	36
Comparison	Beginning	1438.000	86.267	2
	Intermediate	1547.231	143.366	13
	Advanced	1594.300	147.334	20
	Advanced High	1896.301	150.353	39
Total	Beginning	1438.000	86.267	2
	Intermediate	1535.667	137.741	21
	Advanced	1633.255	146.918	51
	Advanced High	1863.240	172.755	75

Appendix J

STAAR 10 Post Hoc Test Years in U.S. Schools

Dependent Variable: STAAR10
Tukey HSD

(I) Years in School	(J) Years in	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
.00	1.00	342.267	130.091	.125	-46.806	731.339
	2.00	289.680	125.421	.247	-86.425	664.784
	3.00	419.195	124.748	.017	46.104	792.287
	4.00	310.983	127.352	.189	-69.897	691.864
	5.00	296.515	126.565	.232	-82.100	675.131
	6.00	294.933	123.742	.213	-75.148	665.015
1 00	.00	-342.267	130.091	.125	-731.339	46.806
	2.00	-52.587	66.693	.986	-252.049	146.875
	3.00	76.923	65.418	.902	-118.722	272.580
	4.00	-31.283	70.257	.999	-241.406	178.840
	5.00	-45.752	68.875	.994	-251.741	160.238
	6.00	-47.333	63.478	.989	-237.181	142.515
2 00	.00	-289.680	125.421	.247	-664.784	85.425
	1.00	52.587	66.693	.988	-146.875	252.049
	3.00	129.516	55.554	.237	-36.632	295.664
	4.00	21.304	61.178	1.000	-161.665	204.273
	5.00	6.836	59.586	1.000	-171.371	185.042
	6.00	5.254	53.255	1.000	-154.020	165.528
3 00	.00	-419.195	124.748	.017	-792.287	-46.104
	1.00	-76.923	65.418	.902	-272.580	118.722
	2.00	-129.516	55.554	.237	-295.664	36.632
	4.00	-108.212	59.786	.544	-287.019	70.595
	5.00	-122.680	58.156	.353	-296.610	51.250
	6.00	-124.262	1.651	.204	-278.737	30.213
4.00	.00	-310.983	127.352	.189	-691.864	69.897
	1.00	31.283	70.257	.999	-178.840	241.406
	2.00	-21.304	61.178	1.000	-204.273	161.665
	3.00	108.212	59.786	.544	-70.595	287.019
	5.00	-14.4682	63.550	1.000	-204.532	175.595
	6.00	-16.050	57.657	1.000	-188.488	156.388
5.00	.00	-296.515	126.595	.232	-675.131	82.100
	1.00	45.752	68.875	.994	-160.238	251.741
	2.00	-6.836	59.586	1.00	-185.042	171.371
	3.00	122.680	58.156	.353	-51.250	296.610
	4.00	14.468	63.550	1.00	-175.595	204.532
	6.00	-1.582	55.964	1.00	-168.958	165.794
6.00	.00	123.742	123.742	.213	-665.015	7.148
	1.00	63.478	63.478	.989	-142.515	237.151
	2.00	53.255	52.255	1.00	-164.528	154.020
	3.00	61.651	51.651	.204	-30.213	278.737
	4.00	57.657	57.657	1.00	-156.388	188.488
	5.00	55.964	55.964	1.00	-165.794	168.958