THE EFFECTS OF TWO CURRICULUM INTERVENTIONS ON THE ACADEMIC ACHIEVEMENT OF ENGLISH LANGUAGE LEARNERS

A Dissertation

Presented to

The Faculty of the Curry School of Education

University of Virginia

In Partial Fulfillment

Of the Requirements for the Degree

Doctor of Education

---

by

Sarah López Rhame, BA, MT

December, 2012

Abstract

Advisor: James Esposito

English language learners (ELLs) comprise a rapidly growing subgroup within the K-12 student population and schools struggle to find instructional programs that can effectively teach both English and academic content. The study explored the differences in academic achievement for students enrolled in two popular and successful instructional programs, an English-only (ESL) or a dual language immersion (DL) program. The purpose of the study was to determine the extent to which pre-test proficiency level, gender, race, and identification as a student with disabilities (SWD) accounted for the variability in academic achievement for students enrolled in an ESL or a DL instructional program. Stepwise multiple linear regression analyses were used to determine the extent of variability in academic achievement accounted for by pre-test proficiency level, gender, race, and identification as student with disabilities (SWD) for students enrolled in an ESL or DL instructional program in five academic areas; overall proficiency, reading proficiency, writing proficiency, listening proficiency, and speaking proficiency. Regression analyses revealed that SWD was a significant variable accounting for the largest percentage of variability no matter the instructional program. The data also demonstrated that the independent variables accounted for more variability in achievement for students enrolled in a DL program versus students enrolled in an ESL program. Further analysis revealed that ESL students had greater gains in proficiency than DL students in three out of the five academic areas studied. Implications for practitioners and further research based on the findings are also presented in the study.

Department of Leadership, Foundations, and Policy
Curry School of Education
University of Virginia
Charlottesville, Virginia

APPROVAL OF THE DISSERTATION

This dissertation, "Factors Affecting Academic Achievement of English Language Learners Enrolled in two Curriculum Interventions", has been approved by the Graduate Faculty of the Curry School of Education in partial fulfillment of the requirements for the degree of Doctor of Education.

Dr. James Esposito, Advisor

Dr. Elizabeth Potts, Committee Member

Dr. Daniel Driscoll, Committee Member

July 16, 2012 Date

ACKNOWLEDGEMENTS

There are many people I would like thank for helping me get to the finish line. I would like to first thank my family for putting up with the late nights, cancelled plans, and postponed vacations. Scott Rhame, my husband, took over household duties when I needed to attend a class or complete an assignment and supported me unconditionally until the end. Karen Rosenfeld, my mother-in-law, graciously offered to babysit so I could attend evening classes. Without her I would have never been able to complete the program. I also want to thank my son, Cole, who had to put up with a "mami" who couldn't play with him as often as he wanted to play. He will now get his mother back.

I would like to thank my work colleagues who allowed me to work flexible hours when needed, and who managed pretty well without me when I needed to take off to write my dissertation. I would like to especially thank Gray Williams, Lucille Dungan, Elissa Gonzalez, Branden Hart, and Daniel Estrada. Without them I could have never completed this study.

I would also like to thank my advisor, James Esposito. "Espo" pushed hard, ever since my first class with him. He challenged me all the way to the end, and as a result I have a product of which I am very proud. Even though it was painful at times, I learned a lot from this experience and am glad I had Espo on my side. He made me a better scholar and for that I am forever grateful.

I also want to thank my committee members. Beth Potts has been a wonderful minor advisor. I hope my APA formatting was up to her exacting standards. Dan Driscoll

graciously accepted to be on my committee when another member dropped out and I cannot thank him enough. I also want to thank all the professors in the program. I use a lot of what I learned in my professional life and hope to one day go back into the public schools and use the tools I have been given to help students and teachers succeed.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS iv

TABLE OF CONTENTS vi

LIST OF TABLES ix

LIST OF FIGURES xii

CHAPTER 1: Introduction 1

Statement of the Problem 6

Purpose of the Study 9

Rationale 11

Limitations of the Study 12

Definition of Terms 12

CHAPTER 2: Review of the Literature 14

History of Teaching ELLs in the U.S. 14

The Gap between ELL and Non-ELL Academic Achievement 16

Factors Contributing to the Gap between ELL and Non-ELL Academic Achievement 18

Instructional Methods for Teaching ELLs 20

English-only instructional programs 23

Dual language immersion programs 25

Comparison of Models in Terms of ELL Academic Achievement 29

Effectiveness of DL programs 29

Effectiveness of ESL programs 32

Large scale studies 35

CHAPTER 3: Methodology 44

Purpose of the Study 44

Research Questions 44

Research Design 45

Treatment 46

Treatment Validity 53

Population and Sample 53

Sources of Data 58

Data Analysis 62

CHAPTER 4: Presentation and Analysis of the Data 63

Descriptive Statistics for Independent and Dependent Variables 63

Analyses to Answer Research Questions 69

Further Analysis 89

Summary of the Data 91

CHAPTER 5: Findings, Discussion, and Implications 93

Summary of the Problem 93

Purpose of the Study 94

Methodology 94

Summary of the Findings 96

Discussion 100

Implications for Practitioners 104

Implications for Further Study 106

Conclusion 109

LIST OF TABLES

Tables	Page
1. Independent Samples T-Test Comparing Continuous Variables between ESL and DL Subgroups	55
2. Independent Samples T-Test Comparing Nominal Variables between ESL and DL Subgroups	57
3. Descriptive Statistics for Total Sample Independent and Dependent Variables	64
4. Descriptive Statistics for ESL Independent and Dependent Variables	66
5. Descriptive Statistics for DL Independent and Dependent Variables	68
6. Correlation Matrix for Overall Academic Proficiency Level Using Pearson Correlation (ESL)	70
7. Correlation Matrix for Overall Academic Proficiency Level Using Pearson Correlation (DL)	71
8. Stepwise Multiple Regression Analysis of Effects of Pre-test Proficiency Level, Gender, Race, and SWD on Overall Achievement for Students Enrolled in an ESL Program	72
9. Stepwise Multiple Regression Analysis of Effects of Pre-test Proficiency Level, Gender, Race, and SWD on Overall Achievement for Students Enrolled in an DL Program	73
10. Correlation Matrix for Reading Academic Proficiency Level Using Pearson Correlation (ESL)	74
11. Correlation Matrix for Reading Academic Proficiency Level Using Pearson Correlation (DL)	75
12. Stepwise Multiple Regression Analysis of Effects of Pre-test Proficiency Level,	

Gender, Race, and SWD on Reading Achievement for Students Enrolled in an ESL Program 76

13. Stepwise Multiple Regression Analysis of Effects of Pre-test Proficiency Level, Gender, Race, and SWD on Reading Achievement for Students Enrolled in an DL Program 77

14. Correlation Matrix for Writing Academic Proficiency Level Using Pearson Correlation (ESL) 78

15. Correlation Matrix for Writing Academic Proficiency Level Using Pearson Correlation (DL) 79

16. Stepwise Multiple Regression Analysis of Effects of Pre-test Proficiency Level, Gender, Race, and SWD on Writing Achievement for Students Enrolled in an ESL Program 80

17. Stepwise Multiple Regression Analysis of Effects of Pre-test Proficiency Level, Gender, Race, and SWD on Writing Achievement for Students Enrolled in a DL Program (N=75) 81

18. Correlation Matrix for Listening Academic Proficiency Level Using Pearson Correlation (ESL) 82

19. Correlation Matrix for Listening Academic Proficiency Level Using Pearson Correlation 83

20. Stepwise Multiple Regression Analysis of Effects of Pre-test Proficiency Level, Gender, Race, and SWD on Listening Achievement for Students Enrolled in an ESL Program 84

21. Stepwise Multiple Regression Analysis of Effects of Pre-test Proficiency Level, Gender, Race, and SWD on Listening Achievement for Students Enrolled in a DL Program 85

22. Correlation Matrix for Speaking Academic Proficiency Level Using Pearson Correlation (ESL) 86

23. Correlation Matrix for Speaking Academic Proficiency Level Using Pearson Correlation (DL) 87

24. Stepwise Multiple Regression Analysis of Effects of Pre-test Proficiency Level, Gender, Race, and SWD on Speaking Achievement for Students Enrolled in an ESL Program 88

25. Stepwise Multiple Regression Analysis of Effects of Pre-test Proficiency Level, Gender, Race, and SWD on Speaking Achievement for Students Enrolled in a DL Program 89

26. Total Variances for Academic Areas, Mean Scores, and Mean Gain Scores for ESL and DL Programs 90

LIST OF FIGURES

Figure	Page
1. Language Acquisition for School	22
2. WIDA Performance Definitions	59

CHAPTER 1

Introduction

United States public schools are faced with the pressing challenge of meeting the instructional needs of increasingly diverse multiracial students. In 1980 the majority of the U.S. population was white: Whites 66%, Hispanics 15%, Blacks 12%, and Asia/Pacific Islanders at 2% (Aud, Fox, & KewalRamani, 2010). Over the past 32 years the majority population has been shifting from white to multiracial. The U.S. Education Department 2009-2010 public school membership report illustrates this increasing diversity. White students comprised only 54% of public school students in 2009-2010. The rest of the student membership for that school year was comprised of Latino/Hispanics at 22%, Blacks/African Americans at 16.8%, Asian/Pacific Islanders at 5.1%, American Indian/Alaska Natives at 1.2%, and those identifying themselves as two or more races at 0.7% (Chen, 2011).

With the shift toward a more multiracial U.S. population, the existing language barrier in public school education has become more pronounced. Many students are English language learners (ELLs). For this study, ELLs are defined as students who have a native language other than English and "whose difficulties in speaking, reading, writing, or understanding the English language may prevent the individual from succeeding in the classroom and on state assessments" (Edvantia, 2007, p. 2).

This increased diversity is closely tied to the rise in the population of people living in the U.S. that were born in other countries. In 2007, about 14% of the U.S.

population was born outside the U.S. (Aud et al, 2010). Many of those born outside the U.S. are ELLs. In U.S. public schools, ELLs comprise up to 10% of the school aged population in certain states (NCES, 2011). The national ELL student population has increased 57% since 1995 compared with 4% for all students (NCELA, 2008). In 2007-2008 approximately 5,318,164 ELLs were enrolled in U.S. public schools (NCELA, 2008). Even though Virginia has a small ELL population, just under 7% of the total student population, the ELL population has almost doubled since the passage of No Child Left Behind (NCLB) from 49,840 in 2001 to 86,751 students in 2009 (Virginia Department of Education, 2009). The significant rise in the ELL population presents challenges to schools nationwide including Virginia. There is a need to find quality instructional programs that teach both English and other academic content to ELLs to prepare them for success in mainstream U.S. public schools.

ELLs comprise a rapidly growing subgroup within the K-12 student population and their educational needs have become a pressing issue in public schools (Maxwell, 2009). Viadero (2009) reported that the challenge public schools continue to face with ELLs is resolving how to progress beyond teaching basic language skills to teach comprehension and other academic content effectively so ELL students do not fall further behind. The U.S. Department of Education requires schools to provide instructional programs for ELLs to learn English as part of the Title III program (Elementary and Secondary Education Act of 2001). In the 2007-2008 school year, 4.7 million ELL students received Title III services (Boyle, Taylor, Hurlburt & Songa, 2010).

One reason that ELL achievement is a persistent problem may be the way the group is defined. Most NCLB subgroups are comprised of a permanent demographic

marker, such as race or gender, but the ELL subgroup is transient. As students learn English and move out of the subgroup new ELL students with little or no mastery of English join the subgroup. Thus the ELL group "continually receives new, mostly lower performing students" and the group is at a constant disadvantage (Kim and Herman, 2009, p.213). ELL students receiving language instruction will show progress, but that progress will be aggregated with new students who have not had significant instruction in English or academic content in U.S. schools, showing no ELL subgroup growth.

In addition to the struggle of acquiring language proficiency, ELL students are more likely to qualify for free and reduced lunch, drop out of high school, be retained, and be more transient than other subgroups (Kim, 2011). Kim studied a cohort of ELLs from 2006-2008 and found 74% qualified for free and reduced lunch, 25% did not graduate from high school, 46% were retained at some point in their academic career, and 39% moved to other schools during their education career (2011, p. 10). Menken also found increased dropout rates in the ELL population, especially after increased graduation requirements and high school exit exams were implemented. Menken found that the ELL dropout rate in New York increased 8% once the high school exit exams were instituted for ELLs while there was only a 1% increase for non-ELLs during that same time period (2010, p. 126). Menken's findings for New York are consistent with national data that graduation rates are lower and high school dropout rates higher for ELLs in states where there are high school exit exams (Dee & Jacon, 2006; Warren, Jenkins, & Kulick, 2005, as cited in Menken, 2010).

Because of the increasing underperformance of ELLs, policy makers identified the need to assess how ELLs were performing in public schools and developed mandates

for states to report on the progress of ELLs as part of NCLB. Currently schools must report on the progress of all students—students from major racial and ethnic groups, students with disabilities, economically disadvantaged students, and Limited English Proficient students (USED, 2003). Because ELLs are a subgroup assessed under NCLB, schools are encouraged to find ways to increase ELL achievement as measured by state mandated assessments. As a result, schools need to develop instructional programs for ELLs that focus on both acquiring English proficiency as well as focus on ways to increase academic achievement.

According to NCLB, ELLs must be tested annually for English language proficiency beginning their first year of enrollment in U.S. schools. Also, in their first year of enrollment, ELLs must be tested in math (Elementary and Secondary Education Act of 2001). ELLs have a one year reprieve in reading, and their reading scores do not count for NCLB until the year following their enrollment in U.S. schools (Elementary and Secondary Education Act of 2001). In terms of NCLB, academic achievement for ELLs falls into two major categories: English proficiency and academic proficiency. Under Title III, states must test their ELLs on English language proficiency annually. Each year, the number of ELLs improving in English must increase along with the number of students gaining English proficiency (Elementary and Secondary Education Act of 2001). In addition, states must assess all major subgroups, including ELLs, in math, reading, and science. To fulfill these mandates, states and local school divisions employ programs to teach ELLs both content and English.

With the mandates and accountability requirements set by NCLB, school divisions need to develop programs that provide ELLs with an engaging and rigorous

curriculum that facilitates academic success. The academic performance of ELLs nationwide has been and currently is lower than that of other NCLB subgroups and the gap between ELLs and native English speakers has not decreased significantly over time (Abedi & Dietel, 2004). The national findings are also true in Virginia. Currently there is an achievement gap in mathematics and reading between ELLs and non ELLs in Virginia. According to the 2009 Virginia State Report Card, 83% of ELLs passed reading Standards of Learning tests compared to 89% of all students, and 79% passed mathematics Standards of Learning tests compared to 86% of all students (Virginia Department of Education, 2010.).

In addition to ELLs not meeting reading and mathematics requirements, they also fail to meet English language proficiency requirements. In the 2007-2008 school year, only 11% of the states met their Title III performance goals for ELLs (Boyle et al., 2010). Failure to meet the English language proficiency mandates is an indicator of a much larger problem. If ELLs cannot gain proficiency in English, they will not be able to have the language skills necessary to pass state mandated high stakes testing required by NCLB (Abedi, 2007; Boyle et al., 2010).

In order to help ELLs succeed, the Elementary and Secondary Education Act provides funding for instructional programs under Title III. These instructional programs are believed to be the means by which ELLs attain the necessary knowledge and skills to learn a new language as well as master academic content. These instructional programs must provide sound instruction that leads to student success. Currently there are two main instructional program models for teaching ELLs, programs that focus on developing literacy in two languages and programs that focus on developing literacy using only

English. The effects of these programs should be closely monitored to assess if they are working for ELLs.

Statement of the Problem

Francis and Rivera (2007) determined that part of the challenge of teaching ELLs was their heterogeneity and their need to acquire academic language proficiency to succeed in content classes. The heterogeneity of ELLs is present in areas such as prior schooling, level of proficiency in the native language, cultural origin, socio economic status, and academic outcomes (Francis & Rivera, 2007). These differences sometimes make it difficult to determine the most appropriate program models for these students. ELLs in the NCLB era must not only learn conversational English which allows students to interact with others in their new country, but also learn academic English and content material, and in many cases in one year's time.

Academic English is vital to succeed in a U.S. public school. Academic English can be defined as the language of texts and formal writing and allows students to gain access to content (Coleman & Goldenberg, 2010). Academic language is more cognitively complex, more challenging to learn, and critical for academic achievement (Brisk & Jeffries, 2008; Coleman & Goldenberg, 2010; Cummings 1984). Francis and Rivera state that "individuals who demonstrate effective use of academic language are able to extract meaning of new content, process it, and add it to previous knowledge" (2007, p.16). ELLs must be able to effectively use academic English to succeed in a public school setting. Determining what program can best instruct ELLs in conversational and academic English as well as content specific material is a major challenge facing educators of ELL students.

Instructional programs for ELLs fall into two broad categories: programs that focus on developing literacy in two languages and programs that focus on developing literacy using only English. Dual language immersion (DL) or bilingual education is one major instructional program. This approach focuses on developing literacy in two languages (Office of English Language Acquisition, 2008). DL programs are long term programs, usually 4-6 years, that use English and the native language to teach content and language objectives (Christian, Howard & Loeb, 2000; Collier, 1995; Genesee, Lindholm-Leary, Howard & Christian, 2003; Valverde and Armendáriz, 1999). DL programs provide instruction in each language for a percentage of the school day. The first DL model is the 90:10, which starts with 90% instruction in the native language and 10% instruction in English during early elementary grades and transitions to 50% English and 50% native language by upper elementary. The second DL model is 50:50, and 50% of the instruction in English and 50% of the instruction in the native language throughout the length of the program (Christian et al., 2000).

DL programs implement the major tenets of second language acquisition theory which assumes that students who become highly proficient in their native language become very successful second language learners. Collier (1995) asserts that language minority students who learn academic skills in knowledge in their first language will be better prepared to learn a second language. Krashen (2008) encourages having a lot of print material in the first language available for students so that students can build up their native language skills to transfer them to learning English. Providing a solid foundation in the native language through instruction and immersion in the native language is an integral part of DL instruction.

DL programs have three major instructional components that implement second language acquisition principles into classroom instruction (Howard & Christian, 2002; Lindholm-Leary, 2001).The first component is that the program must include equal numbers of ELLs and English dominant students. This is a critical component that differentiates DL from foreign language immersion because dual language immersion programs have a mix of students who are fluent in both languages used in instruction. DL instructional models have students who are ELLs and students who are native English speakers, so the students have peers that are dominant in a language other than their own native language. The second component is that the program must be integrated so that the ELL and language majority students are together for academic instruction. The third component is that DL programs must provide core academic content to both groups of students in both languages, and by third grade they should be receiving literacy instruction in both languages. Dual language immersion allows teachers to deliver academic instruction in the native language, thus allowing students to learn in their native language. It is assumed that students then transfer the academic concepts to their new language and start making gains academically.

Programs that focus on developing literacy using only English include programs in which students work only on English grammar, vocabulary and communication skills or programs that blend both content and English instruction in an English only environment (Office of English Language Acquisition, 2008). Traditional English as a Second Language (ESL) instruction "…is explicit, direct instruction about the English language intended to help ELLs "catch up" to their student peers who are proficient in English. It includes learning outcomes in speaking, listening comprehension, reading, and

writing English" (Smith, Coggins & Cardoso, 2008. p. 296). ESL programs begin with teaching students social language so students can navigate the new school community. Collier defines social language as language that "…includes the development of basic literacy for the use in situations such as shopping, use of transportation, access to health services, writing a letter to a friend, or sending an email message" (1995, p. 8). To teach students academic content, traditional ESL programs may employ sheltered English instruction where the content is modified to make it more comprehensible (Rossell, 2005; Smith et al., 2008). Teachers using sheltered English instruction to teach content will have both content and language objectives. As students become more proficient in English, they transition out of ESL classes and move into classes that are mainstream.

Traditional ESL and DL programs both address the instructional needs of ELLs. There are strong proponents of the dual language method (Alanís & Rodriguez, 2008; Genesee et al., 2005) as well as the English only method (Echevarria, Short, & Powers, 2006; Rossell, 2005). As a result, school districts may employ more than one instructional program. Students may also move from one instructional program method to another when switching schools or even grades within one school. This variability in instruction may be a cause for the continued lack of progress for ELLs.

Purpose of the Study

The purpose of the study was to determine the extent to which pre-test proficiency level, gender, race, and identification as a student with disabilities (SWD) account for the variability in academic achievement for students enrolled in an ESL instructional program or a DL instructional program. Analyzing both a traditional English only ESL program and a DL program will address one of the main issues cited by researchers: that

ELL treatments are evaluated in isolation without either a control group or a comparison group (Conger, 2010; Ramirez, Yuen, Ramey, Pasta, &Billings, 1991; Rolstad, Mahoney, & Glass, 2005; Rossell & Baker, 1996; Slavin & Cheung, 2005). Using a pre-test/post-test design addresses another gap in the research since most studies use only summative data to compare programs or measure the effects of an instructional program (Conger, 2010; Ramirez, Yuen et al., 1991; Rolstad et al., 2005; Rossell & Baker, 1996; Slavin & Cheung, 2005).

The research questions examined the extent to which the independent variables accounted for variability in overall, reading, writing, listening, and speaking academic achievement for students enrolled in different instructional programs.

1. To what extent does pre-test proficiency level, gender, race, or SWD explain variability of overall academic achievement for students enrolled in an ESL program versus students enrolled in a DL program?
2. To what extent does pre-test proficiency level, gender, race, or SWD explain variability of reading academic achievement for students enrolled in an ESL program versus students enrolled in a DL program?
3. To what extent does pre-test proficiency level, gender, race, or SWD explain variability of writing academic achievement for students enrolled in an ESL program versus students enrolled in a DL program?
4. To what extent does pre-test proficiency level, gender, race, or SWD explain variability of listening academic achievement for students enrolled in an ESL program versus students enrolled in a DL program?

5. To what extent does pre-test proficiency level, gender, race, or SWD explain variability of speaking academic achievement for students enrolled in an ESL program versus students enrolled in a DL program?

Rationale

In this era of accountability, school divisions need to find a way to teach both content and language to ELLs. Callahan, Wilkinson, and Muller state, "Understanding how to best serve ELL students in U.S. schools requires… an evaluation of the treatment and achievement of students once identified for services." (2010, p. 108). This study examines the effects pre-test proficiency level, gender, race, and SWD have on the achievement of students enrolled in a DL or ESL instructional programs to determine what program has more of an impact on student achievement.

Study findings provide an objective way to help determine high quality instructional programs schools choose for their ELL populations (Calderón, Slavin, & Sánchez, 2011; Conger, 2010). The findings of this study may aid school divisions with limited resources to target those resources towards programs that demonstrate the greatest academic achievement. School divisions that implement both programs can use the data to assess if there is a need for both programs or if one program can address the needs of ELLs in their division.

The findings may also inform staff development in the area of ELL instruction. Teacher training can focus on the more successful model to save fiscal resources. Teachers may review the strategies from either DL or ESL that make the most impact on student achievement and implement those to teach ELLs. Short, Echeverria, and Richards Tutor (2009) advocate for professional development and sustained teacher training due to

promising research findings, and the findings from this study could be used by schools to justify training for either DL or ESL instructional methods. There are several ways these data can be used to inform the issue of ELL student achievement.

Limitations of the Study

The study has several limitations. One limitation deals with the inability to assess the fidelity of the implementation. Linked to fidelity is the teachers' level of competency. Teachers who do not have the capability to implement the model as intended can greatly affect the outcomes.

This study is limited to the definition of academic achievement. For this study academic achievement is defined as a score on the Assessing Comprehension and Communication in English State to State for English language learners (ACCESS for ELLs). The ACCESS is one of many assessments of academic achievement.

The methodology of the study only focuses on the independent variables of pre-test proficiency level (overall, reading, writing, listening, speaking), gender, race, and SWD. The study does not control for other variables that may impact academic achievement.

Definition of Terms

Achievement: For this study, achievement will be measured by proficiency level scores on the ACCESS. Students receive an overall score on English language proficiency that is comprised of the mean of four sub scores in reading, writing, listening, and speaking.

Treatment: For this study, treatment is defined as the instructional program in which ELLs are enrolled. The two programs are traditional ESL and DL.

Pre-test: For this study, the pre-test is defined as the Fall ACCESS pre-test proficiency level scores in overall proficiency, reading proficiency, writing proficiency, listening proficiency, and speaking proficiency.

Post-test: For this study, the post-tests are defined as the Spring ACCESS scores post-test proficiency level scores in overall proficiency, reading proficiency, writing proficiency, listening proficiency, and speaking proficiency.

CHAPTER TWO

Review of the Literature

There are several areas of research that are important to the study: history of teaching ELLs in the United States, the gap between ELL and non-ELL academic achievement, factors contributing to the gap between ELL and non-ELL academic achievement, the instructional models for teaching ELLs, in particular ESL and DL models, and the comparisons of those models in terms of ELL academic achievement.

History of Teaching ELLs in the U.S.

Teaching ELLs has been a part of U.S. education since the early days of the country. Bilingual education programs were part of public and private schools from the 1700s -1900s. In areas with large immigrant populations there were private, public, and parochial schools that provided education for ELLs using English-only and bilingual education (Baker, 2001). In the early 20th century attitudes toward using bilingual education as a way to teach ELLs changed. Federal legislation passed in 1919 required elementary students to receive instruction solely in English in all public and private schools (Baker, 2001).

Bilingual education and DL programs regained popularity in the 1960s as a result of the influx of Cuban immigrants to Florida. In 1963 Cuban exiles opened Coral Way Elementary School, a dual language immersion school in Dade County, Florida (Baker, 2001). Cuban immigrants wanted to preserve their native language while awaiting a safe time to return to Cuba. Florida government officials provided support and funding to the

school because they felt this was a short term experiment and it was a way to support anti-communist causes (Baker, 2001). In 1968 bilingual education received an additional boost with the passage of Title VII of the Elementary and Secondary Education Act that authorized federal funds for teaching ELLs and for transitional bilingual education programs.

The debate on how to teach ELLs in recent history began with the *Lau vs. Nichols* case in 1974. The *Lau* case focused on Chinese immigrant students that received education only in English without any support to learn the new language or academic curriculum. The Supreme Court mandated that schools had to provide full access to education for ELLs but did not specify a particular method of instruction (as cited in Edvantia, 2007). Throughout the country Lau remedies were enacted, such as English as a Second Language classes, English tutoring, and some bilingual education programs (Baker, 2001). In 1978 Congress amended the Bilingual Education Act and restricted federal funding to only support transitional bilingual education programs for ELLs (Baker 2001; Edvantia, 2007). As a result, programs that delivered English-only instruction or had a short bilingual component became the standard programs in most schools since no federal funding could be used to maintenance bilingual education programs.

The debate intensified where transitional bilingual programs were used. One group advocated that students needed to achieve high levels of proficiency in their native language to better learn English while another group advocated for teaching English as early and as quickly as possible and assimilating these students into mainstream American society (Baker, 2001).

Proponents of English-only programs gained momentum when the 1984 and 1988 amendments to Title VII increased funding for English-only programs. The Reagan administration also promoted English-only programs by redirecting 25% of federal funding under Secretary of Education William Bennett for English-only programs, such as Sheltered English instruction (Baker, 2001). The Improving America's Schools Act of 1994 brought a closer look at the academic achievement of ELLs and found ways to teach these students that would yield better results.

In the era of standards and high stakes testing brought on by NCLB, focus is shifting to finding programs that increase the achievement of all students, no matter the subgroups to which they belong. Schools are encouraged to find ways to improve their DL and ESL programs to increase ELL achievement as measured by state standards. ELL programs must expand their focus of acquiring English proficiency to include securing ways to increase overall academic achievement.

The Gap between ELL and Non-ELL Academic Achievement

The achievement gap between ELL and non-ELL students is of great concern as the ELL population continues to grow. Data collected by the census bureau in 2009 show that 37% of the total 4th grade and 21% of the 8th grade total student population are Hispanic, and a significant number of them are also ELL (Hemphill & Vanneman, 2011). Fry (2007) analyzed National Assessment of Educational Progress (NAEP) scores and found that 47% of 4th grade ELLs and 51% of 8th grade ELLs scored lower than their native English speaking counterparts in reading. The math gaps were not much better with 35% of 4th grade ELLs and 51% of 8th grade ELLs scoring lower than their English speaking counterparts. NAEP results from 2005 found that 46% of 4th grade ELLs scores

below basic in math and 73% scored below basic in reading. For 8th grade ELLs the scores are much worse. Seventy-one % of 8th grade ELLs scored below basic in math and reading (Fry, 2007).

Hemphill and Vanneman, 2011, found that while Hispanic students' math scores on the NEAP increased from 1999-2009, the gap in 4th and 8th grades remained significant. In 2009, there was a 21-point gap in 4th grade and a 26 point gap in 8th grade math scores between Hispanic and White students. When analyzing ELL scores there was a 19 point gap in both 4th and 8th grade math scores between ELLs and non-ELLs. Hemphill and Vanneman found the same pattern for reading scores. While Hispanic and White students' scores increased in 2009, there was still a 26-point gap in 4th grade and a 24-point gap in 8th grade between the two groups. When analyzing ELL scores there was a 29-point gap in 4th grade and a 15-point gap in 8th grade between ELLs and non-ELLs (Hemphill & Vanneman, 2011). Kim (2011) found an average 20-point achievement gap between ELLs and non-ELLs in reading for grades 5, 8, and 10 and a 15-point gap on high stakes state math assessment scores in her study of three southwest and western states. Menken (2008) found a gap of 20-50 points nationally between ELLs and non-ELLs on state assessments for English language arts and other content areas.

The academic achievement of ELLs lags behind their native English speaking peers, and ELLs consistently underperform in academic settings (Echeverria et al., 2006). This achievement gap becomes quite apparent when looking at the performance of ELLs in states where high stakes testing make up part of the graduation requirements. Snow and Biancarosa (2003) found there is a large percentage of ELLs that met all graduation requirements but did not receive diplomas because they failed their state's high stakes

tests. Menken (2008) confirmed that finding and found that graduation rates fell for ELLs in California and New York because they failed the high school exit exams. Finding more effective instructional models for ELLs that address the need to learn content as well become proficient in English is one way to help narrow this academic gap.

Factors Contributing to the Gap between ELL and Non-ELL Academic Achievement

An important factor contributing to ELL underachievement is the lack of strong literacy skills in English exhibited by ELLs. Learning conversational English is not enough for ELLs to be successful in mainstream classrooms (Collier, 1995; Echevarria et al., 2006; Genesee et al., 2005; Rossell, 2005). August and Shanahan (2006) found that ELLs and their native English speaking peers had equal literacy development in the areas of decoding and spelling, but when measuring text level skills, such as reading and writing, ELLs fall behind. August and Shanahan (2006) state that the literature on ELL literacy development is small but there are studies that can help frame potential interventions to help ELLs narrow the achievement gap.

Another contributing factor may be the state and federally mandated assessments that are used to the report the achievement gaps between ELLs and non-ELLs. In most states, ELLs take the same assessments as their native English speaking peers but may not have the necessary language skills to understand the assessment, even with accommodations. Like students with disabilities, ELLs face accessibility challenges when taking high-stakes assessments. Unfortunately, many states and school districts provide ELLs with accommodations for students with disabilities since they are not sure which accommodations are appropriate for ELLs. The use of test accommodations originated

with students with disabilities, and therefore, many of the accommodations used for ELLs were created for students with disabilities and may not be appropriate for ELLs (Abedi, 2005). Minnema, Thurlow, Anderson, and Stone (2005) found that several states use the same accommodations for students with disabilities for ELLs. Minnema et al. argue against this practice because, "in general, accommodations developed for students with disabilities do not address the linguistic difficulties faced by ELLs" (2005, p. 8). Rivera (2003) analyzed 73 accommodations used for ELL students on high-stakes assessment and only 11-15% were appropriate for ELLs (as cited in Abedi, 2009, p.22). When analyzing gap scores, one needs to account for the lack of accessibility ELLs face when taking a large scale assessment designed for mainstream students and how that may impact ELL scores.

Another important factor contributing to underachievement is that some ELLs also have learning disabilities that prevent them from being successful. Learning disabilities are difficult to identify in ELLs due to the lack of language proficiency. Schools struggle to assess whether underachievement is due to a lack of English proficiency or a learning disability (Abedi, 2009; Anderson, Minnema, Thurlow, & Hall-Lande, 2005; Artiles, Rueda, Salazar, & Higareda, 2005). Once an ELL is identified as a student with disabilities there are challenges as to how best to serve the student's special needs while providing an English language learning program that will allow the student to be successful in mainstream English classes (Anderson et al., 2005). A policymaker described the challenges of identifying ELLs: "One is identifying them truly as students with disabilities versus students with LEP that may be what appears to make them eligible for special ed. Then there are issues with assessing them not for eligibility but for

achievement." (Anderson et al., 2005, p. 6). The difficulty with identifying ELLs who may also have special needs can lead to overrepresentation of ELLs in special education programs. This overrepresentation can result in ELLs not receiving the appropriate instruction to improve academic achievement.

Artiles et al. (2005) found overrepresentation in several districts in California due to a combination of misidentification and lack of programs for ELLs. This was especially the case in large school districts with scarce resources (Artiles et al., 2005). In their study of California districts Artiles et al. found that ELLs in elementary school who had limited proficiency in both their native language and English were 40-50% more likely to be placed in a special education program than their White peers (2005, p.293). They also found that ELLs in English-only programs were more likely to be placed in special education programs when compared to their peers in DL programs. This is an interesting finding since English-only programs provide little if any first language support. Without the ability to engage these students in their native language, teachers may misidentify a student as special needs when in reality it is the lack of English proficiency that is causing the poor performance. Sáenz, Fuchs, and Fuchs cite the lack of research on "effective teaching strategies for ELL with learning disabilities" as an issue that needs to be addressed to minimize the achievement gap (2005, p. 232). They advocate the use of peer tutors and cooperative learning to help ELLs in the classroom. Both of these instructional methods are prevalent in DL programs.

Instructional Models for Teaching ELLs

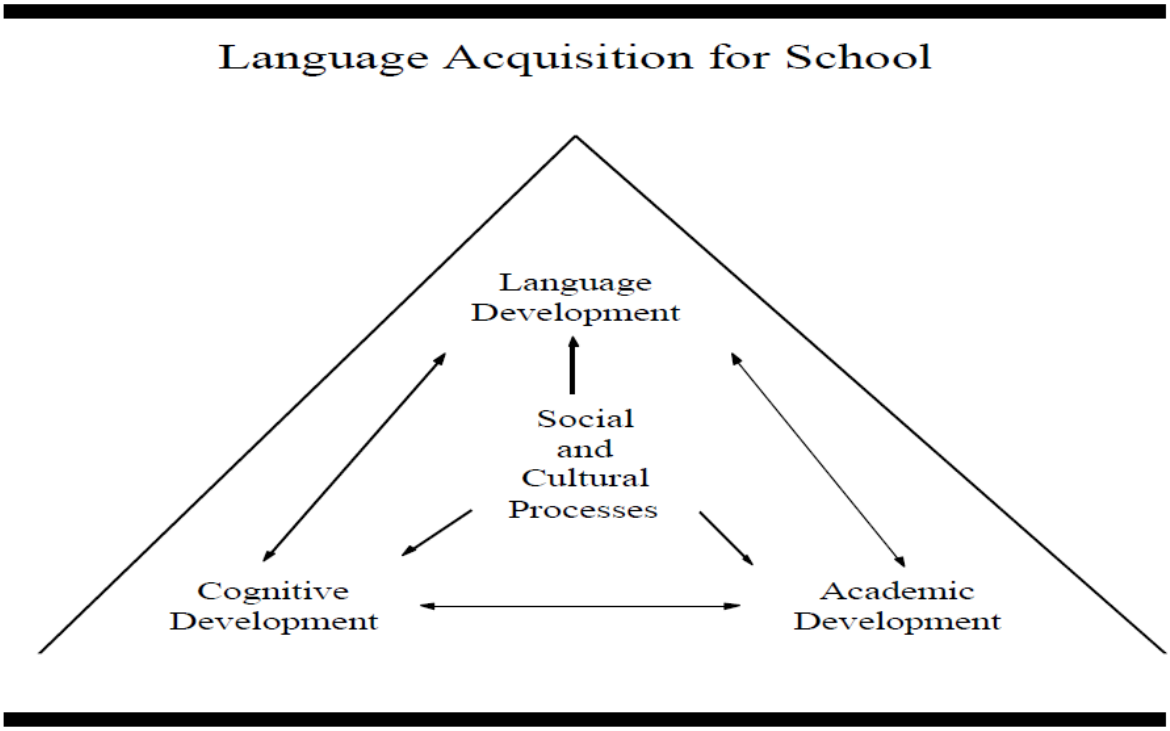
Providing instructional models where ELLs can learn English and other academic content is an important step to help these students succeed. ELLs enrolled in an

instructional program to teach English proficiency, no matter what the program design, benefit greatly academically (Kim & Herman, 2009). Former ELL students reclassified as proficient who were enrolled in a program to learn English outperformed their peers who were not in structured English programs, "…regardless of content areas and grade levels." (Kim & Herman, 2009, p.224)

ELLs enrolled in a language instruction program early perform better academically than their counterparts who are identified later (Collier, 1995; Genesee et al., 2006; Kim & Herman, 2009; Lindholm-Leary, 2001). Kim and Herman found that students who were classified as English proficient in 4th grade "…will likely differ substantially in various characteristics from students who are reclassified in the eighth grade" (2009, p.225). Kim also found in his study of an ELL cohort from 2006-2008 that ELLs who were reclassified as English proficient earlier have more "favorable outcomes" than students who are reclassified later and compared to high school ELLs (2011, p. 11). Kim and Herman's research highlights the need for ELLs to be enrolled in an English language instructional program to be successful in U.S. schools.

Collier's work is often cited by other researchers and used by school divisions to develop English language instructional programs. Like Kim and Herman, Collier advocates for early intervention and provides a conceptual framework widely used by school divisions, including the one in this study, outlining the process of language acquisition. Figure 1 outlines Collier's conceptual framework for language acquisition.

Figure 1: Language Acquisition for School



Adapted from Promoting Academic Success for ESL Students, by V. Collier, 1995, p.21. Copyright 1995 by New Jersey Teachers of English to Speakers of Other Languages – Bilingual Educators.

At the center of Collier's conceptual framework are the social and cultural processes the students are exposed to every day at home and school. The academic, language, and cognitive development parts of the pyramid need to be developed in concert so that a student can receive a well rounded education that allows for the acquisition of language, academic content, and higher order cognitive skills needed to be successful in mainstream classes (Collier, 1995). Collier advocates for the development of both the first and second language "to a deep level of proficiency, for maximum cognitive growth" (1995. p. 22) and cites DL programs as the most effective way to achieve this

deep proficiency because DL programs "focus on the meaningful use of the two languages through academic content across the curriculum" (1995, p. 35).

No matter which program school divisions implement, there are several important elements to effective instruction for ELL students, and divisions should strive to have these present in their ELL instructional models. One element is direct, explicit, and long-term vocabulary instruction (Calderón et al., 2011; Carlo et al., 2008). Calderón et al. define explicit vocabulary instruction as providing exposure to the word in multiple forms, providing examples in several contexts, teaching proper pronunciation and spelling, "and when possible, teaching its cognates, or a false cognate, in the child's primary language" (2011, p.110). To help ELLs with reading skills, content classes should incorporate vocabulary knowledge, reading and writing instruction (Calderón et al., 2011; Echeverria et al., 2006; Lindholm-Leary, 2001; Short et al., 2011). ELL language instruction needs to be "divided among work meaning, decoding, grammatical structures, background knowledge, and comprehension skills" (Calderón et al., 2011, p.111) in order to provide students opportunities to practice the language as well as learn the content. As for instructional methods, cooperative learning has been shown an effective teaching tool in several studies (Calderón, Hertz-Lazarowitz, & Slavin 1998; Carlo et al., 2005; Saunders & Goldenberg, 1996).

English-only instructional programs. Echevarria et al. (2006) advocate English-only sheltered instruction programs as a way to teach students English and provide them a program to access grade-level academic content. The researchers outline techniques that teachers can use, such as "slower speech, clear enunciation, use of visuals and demonstrations, scaffolded instruction, targeted vocabulary development, connections to

student experiences, student-to-student interaction, adaptation of materials, and use of supplementary materials" (Echevarria et al., 2006, p. 199) to provide a rich, English-only environment where students will become proficient in academic English. However, teachers need to be respectful of a students' native language, and not minimize the importance of a student's native language when instructing in an English-only model (Calderón et al., 2011).

Sheltered instruction strategies or Specifically Designed Academic Instruction in English (SDAIE) are traditional English-only instructional methods that help ELLs attain academic English, as well as content knowledge. SDAIE classes use comprehensible input, which means adding visual cues, clarification, and background knowledge to make an academic concept comprehensible to ELLs (Coleman & Goldenberg, 2010; Krashen, 2008). In the Sheltered Instruction Observation Protocol (SIOP), an English-only instructional method, ELLs can learn academic content in a sheltered environment along with specific language objectives that help students attain academic English (Coleman & Goldenberg, 2010; Echeverria et al., 2006, Genesee et al., 2006, Hansen-Thomas, 2008). In these sheltered environments, students can engage in content-specific lessons that provide explicit English language instruction to increase their content knowledge and English language proficiency. Effective sheltered instructional programs provide ELLs with high quality content instruction, similar to the instruction native English speakers receive, along with explicit instruction in academic English. Sheltered English instruction does borrow from DL in that it allows the native language to be used when appropriate, such as to clarify academic content (Hansen-Thomas, 2008).

Baker (1998) defines structured English immersion (SEI) as a program where English is used and taught at a level appropriate to students, and where teachers strive to provide English-only instruction as much as possible. Baker proposes small classes, like the one in the sample used for this study, so that there can be ample time for dialog between teacher and students. Baker conducted his study in El Paso, where there were well established English-only programs, and where students in a SEI program outscored bilingual education students for 11 straight years. According to Baker, the essential component to a successful SEI program is adjusting the English instruction to the level of the ELL. Baker says when advocating for SEI,

> There seem to be only two ways to screw up a LEP child's opportunity to learn English. The first is to use too much of the non-English language in the classroom, and this seems to be the problem in many bilingual education programs. The second is to fail to realize that LEP students face a more demanding task in school than do native English-speaking students. LEP students have to learn everything in the curriculum and then learn English on top of it (1998, para. 26).

Dual language immersion programs. Christian et al. define dual language immersion programs as "programs that integrate language minority and majority students in the same classroom, beginning in the early elementary grades, and provide content area instruction and language development for all students in two languages. [DL] programs seek an environment that promotes positive attitudes toward both languages and cultures and supports the development of full bilingual proficiency for both groups of students" (2000, para. 2)

Dual language is one approach to bilingual education where the primary goal is to develop fluency in two languages, where one language is the student's native language and the other is a second language. The development of the native language is an important tenet of DL programs since second language acquisition research states that the more developed the first language, the easier it is to learn the second language (Cummings, 1992; Krashen, 2008). Ramirez, Yuen, Ramey, Pasta, and Billings found this to be the case in their large scale longitudinal study where students in DL programs with stronger first language skills learned English faster than those with less proficiency in their native language (1991, p. 510). To achieve this goal of bilingualism, DL programs should be at least six years, k-5 (Collier & Thomas, 1997; De Jesús, 2008). While some research has shown that any exposure to a ELLs native language positively impacts achievement, many transitional or maintenance programs do not provide any better results than traditional English-only ESL instruction (Cobb, Vega, & Kronauge, 2006; Rossell & Baker, 1996; Thomas & Collier, 1997).

DL programs provide academic instruction to ELLs and non-ELLs in two languages, where one is the primary language of each group of students. The goal of DL programs is to develop bilingualism, biliteracy, and academic achievement (Lindholm-Leary & Block, 2010). The instruction in both English and another language has been shown to have a positive correlation in achievement in both English and the second language (Cummings 1992, Lindholm-Leary, 2001).

Thomas and Collier (1997, 2001) conducted longitudinal studies that showed the effectiveness of bilingual education and found that well implemented DL programs provide ELLs with the best opportunity to succeed in mainstream English classrooms.

Thomas and Collier also found that that it takes four to seven years for students in DL programs to outperform other ELLs, but the fact that these students are learning both their native language and English, makes their English proficiency much higher than ELLs in traditional ESL classrooms (Thomas & Collier, 2001). The work of Thomas and Collier is widely cited by researchers analyzing effective ESL instruction and their findings are used by school divisions when justifying their ELL instructional programs.

Lindholm-Leary and Block (2010) studied the effectiveness of 90/10 DL programs in California and used the summative California Standards Test (CST) to measure achievement in math and reading for ELL students in these programs. They also measured Spanish math and reading achievement by using the Aprenda, a national norm referenced test written in Spanish. Lindholm-Leary and Block (2010) found that ELLs that had just entered the program scored slightly higher in reading than their ELL peers in English-only programs and, by fifth grade, the ELLs in DL programs were outscoring their peers in English-only programs by 13% (2010, p. 51). They also found that these ELLs were performing at the same rate as their English proficient Hispanic counterparts. Lindholm-Leary and Block also analyzed math scores and found similar results. ELLs who had just entered the program performed slightly lower, four percentage points, than their ELL counterparts in English-only programs. However, by the end of elementary school, ELLs in DL programs outscored their peers in English-only programs 67% to 25% (2010, p. 52). Lindholm-Leary and Block (2010) found that students in DL programs were not only more successful that their peers in English-only programs but narrowed the achievement gap between them and their native English speaking peers.

Cobb et al. (2006) conducted a study to measure the effectiveness of elementary DL programs on ELL middle school academic achievement. The students in the study were in a 50/50 model where instruction is evenly divided between English and Spanish throughout the instructional day. The researchers used the Colorado Student Assessment Program (CSAP) summative measure to analyze reading, writing, and math scores. Cobb et al. found that DL programs were as effective as if not more effective than traditional ESL programs, especially in the areas of reading and writing. ELL students in the DL programs outperformed students in English-only program in all three CSAP tests (2006, p. 39). In addition, the researchers found that DL programs were beneficial to native English speakers as well as may help allay fears of parents and policymakers who worry native English students may lose ground in other content areas while learning a second language.

Genesee et al. advocate dual language immersion programs as the way for ELLs to succeed in mainstream schools. ELLs who have a strong native language become successful bilinguals who deploy bilingual strategies to figure out cognates between the two languages, use judicious translation, and use prior knowledge in the native language to figure out unfamiliar material in English (Genesee et al., 2005). Tong, Irby, Lara-Alecio, and Mathes (2008) conducted a three year randomized study on the effectiveness of DL programs and found that with high quality instruction that included phonemic awareness, letter-sound correspondence, word recognition, spelling, fluency, and comprehension, ELLs shows great gains in English.

Comparison of Models in Terms of ELL Academic Achievement

Effectiveness of DL programs. There are research studies that demonstrate the success of both instructional methods. Lindholm-Leary, one of the prominent researchers on DL, has found that ELLs in dual language programs are more successful in developing English language proficiency, passing high school exit exams, and reading and math achievement than their counterparts in traditional ESL programs (Lindholm-Leary & Genesee 2010; Lindholm-Leary & Howard, 2008). Studies have also shown that ELLs in DL programs have higher achievement scores, grades, and achievement outcomes than their native English speaking peers (Genesee et al., 2006; Lindholm-Leary, 2001). In 2011, Lindholm-Leary and Hernandez conducted a study on DL student achievement using the California Standards Test (CST), a summative assessment, and disaggregated for ELLs and those reclassified as English proficient. ELLs who went through a DL program and were reclassified as English proficient outscored their native English speaking peers and confirm other studies that show that bilingual students have higher achievement levels than their English-only peers (Genesee et al., 2006; Lindholm-Leary, 2001; Lindholm-Leary & Hernandez, 2011). Lindholm-Leary and Hernandez also found that 85% of the ELLs in their study were on their way to achieving English language proficiency and cite this as an indication that the "additive bilingualism and biliteracy" (2011, p. 542) of DL programs allow students to achieve success mastering academic English.

Alanís and Rodriguez (2008) analyzed Texas state assessment test (TAKS) scores, a summative measure, in their study on the academic achievement of fifth grade ELLs who received instruction in a DL program. The study centered on a well

established elementary DL program using the 90/10 model. Alanís and Rodriguez (2008) developed their study to answer the questions of participation in the program as a contributing factor to academic achievement and factors that allow for the sustainability of the DL program. The procedures included site visits and observations of DL classrooms and interviews with DL teachers. Alanís and Rodriguez compared DL students to ELLs in English-only programs when conducting analyses of the Texas TAKS assessment scores. The DL students in the study conducted scored consistently higher than their ELL peers that were in English-only programs. Alanís and Rodriguez concluded that the scores for the DL students were "impressive given that students were tested in English even though students in the dual language program did not receive formal English reading until the third grade" (2009, p. 310) and that instruction in Spanish did not negatively impact their English language development.

De Jesús (2008) conducted a study in an urban district with a newly implemented 50/50 immersion program and found that ELLs as well as native English speakers outperformed students in the general population. De Jesús gathered data on 4th grade for seven years after the implementation of the DL program and the data showed "a constant and impressive pattern of student success" (2008, p.206). As with other studies, De Jesús used summative data, a state mandated high stakes test in reading and math, as a measure of achievement. However, De Jesús focused on comparing two instructional methods to find which method was more effective. The district in his study had a school system mandated transitional bilingual education program, a program where ELLs spend 2-3 years learning English with limited use of the native language to clarify concepts and answer questions. By fourth grade, ELL students in the DL program were 60% proficient

while ELL students in the transitional bilingual program mandated by the school system were only 37% proficient as measured by the NCLB required reading test (2008, p. 203). Students in the DL program also showed greater gains in Spanish language proficiency, increasing their proficiency by 7% while the students in the state mandated bilingual program only increased their Spanish proficiency by 3% (2008, p. 204). De Jesús argues that DL programs work because they experience what he calls a "cognitive stretch" where students think inferentially and "become active, higher-order thinkers" (2008, p. 208).

Valverde and Armendáriz conducted a study comparing several instructional methods for ELLs including pull-out, structured immersion, transitional, maintenance, and dual language (1999, p.1). The researchers outline several weaknesses of traditional English-only programs. ESL pull-out programs have several negatives including time out of regular classroom instruction, potential low self esteem from being pulled out while their English-only peers remain in the classroom, and added expense to hire teachers for just the pull-out part of the programs (Valverde & Armendáriz, 1999). Structured immersion programs are weak because even though they allow the use of the native language in certain situations it is not enough to develop literacy. The researchers outline several strengths of DL programs including an emphasis on concept development, academic development in both languages, separation of languages depending on the subject matter taught, and fostering of the home language as a resource (Valverde & Armendáriz, 1999). Valverde and Armendáriz are proponents of DL programs because they effectively develop both language and academics, so that ELLs can be better prepared for mainstream classrooms once they exit the program.

Effectiveness of English-only programs. Rossell, a strong proponent of traditional English-only ESL programs, attributes the perceived success of DL program to the inclusion of many programs that may not qualify as DL. Rossell states that important parts of bilingual education theory, such as the need to have a well defined first language to develop a second language and that development of the second language is greatly benefited by the first language, do not account for the great variation in languages spoken by immigrants to the U.S. (2003). Ideographic languages, such as Japanese and Chinese, are much harder to master than phonetic languages and, Rossell argues, students would be best served to learn English first since it is easier to learn than some of the ideographic languages. Rossell extends her argument to non-Roman alphabets, stating that teaching young children in a language so different from English can confuse young students (2003). Rossell concludes that ELLs in traditional ESL programs gain proficiency faster than ELLs in DL programs (2003, 2005).

Conger (2010) analyzed ELL scores in New York City and found that students in DL programs learn English less quickly. Conger cites the recent votes in California, Arizona, and Massachusetts for English-only ESL instruction to show that DL instruction is not as popular as it once was. Like Rossell, Conger states that there have been many studies on the effectiveness of English language learning programs but "there have been no large-scale experiments and most of the prior studies rely on small samples of students in one or two schools and use single pre- and post-test group designs without comparison between instructional approaches" (2010, p.1106). Conger's study followed cohorts of students from 1996-2002 and evaluated three treatments: students receiving native-language instruction (DL), English-only ESL, and no English language services (2010,

p.1109). In Conger's group students who had lower English proficiency tended to go into DL programs while students with higher English proficiency went into traditional ESL programs. Conger found that after the first year 90% of students in the DL programs were still classified as ELL while only 66% of students in traditional ESL programs were still classified as ELL; and four years later 52% of DL students were classified as ELL while only 23% in the ESL programs were ELL (2010, p. 1112). The percentage after four years is significant because proponents of DL programs argue that it takes 4-6 years to learn English proficiently in these programs (Collier, 1995; Genesee et al., 2005; Lindholm-Leary, 200l; Thomas & Collier, 1997). Conger (2010) also found that students who tended to be older, poorer, diagnosed with disabilities, and who frequently transferred schools were less likely to become proficient in English. Conger compared students who were enrolled in bilingual education programs and who were Spanish speaking, Non-Spanish speaking, and Chinese speaking and found that all three groups took longer to become English proficient but did not correlate the outcome to the native language instruction. Conger (2009) concluded that it was negative selection into those programs rather than the programs themselves that were the cause of the decreased in ELL proficiency. Conger found that DL students in his study never caught up to their ESL counterparts. Conger's conclusion is that bilingual education "either interferes with English-language acquisition or has no effect" (2010, p. 1119). Conger recommends more random studies with larger populations to see which instructional program is best for ELLs because of the large amounts of money being spent by federal, state, and local governments on bilingual education. Conger argues that there is not a "clear positive effect of bilingual education" in his study and that "raises the question of whether federal,

state, and local governments should invest further resources in a program that may not be more effective than ESL pullout services in helping students acquire English proficiency" (2010, p. 1119). Conger does state that if bilingual education has other benefits, such as the learning of academic concepts more efficiently in the native language, then it should be further studied, but should not be labeled as a better way to learn English.

Short et al. analyzed the effectiveness of English-only instruction, specifically sheltered English instruction. Short et al. (2011) define sheltered English instruction as subject level instruction provided in English in a classroom where most if not all students are ELLs. The Sheltered Instruction Observation Protocol (SIOP) intervention is an approach where ELL teachers "teach subject area curriculum to students learning through a second language using techniques that make the content material accessible and also helps develop the students' second language skills" (2011, p. 364). The SIOP model is an "approach for integrating language and content instruction in either content areas or language development classes" (2011, p. 364). Short et al. conducted an ANOVA to see if students instructed in sheltered English classrooms where teachers used the SIOP protocol performed better in reading, writing, and oral proficiency than a comparison group. Students instructed in sheltered English classrooms where teachers used the SIOP protocol scored 33% higher in oral language, 28% higher in reading, 36% higher in writing and had a total proficiency level 77% higher than the comparison group (Short et al.,2011, p.370). Short et al. (2011) concluded that sheltered English instruction was an effective way to increase proficiency in oral language, reading, and writing.

Callahan et al. (2010) analyzed the effectiveness of ESL placement in terms of future success throughout high school and the ability to prep for college. They argue that

"if ESL placement constrains students' exposure to academic content, the cumulative effects may be substantive and significant" (2010, p.87). Furthermore, they state that if ELLs have less challenging coursework it will negatively affect their English language acquisition and their ability to integrate into mainstream society (Callahan, Wilkinson, & Muller, 2010). Callahan, Wilkinson, and Muller hypothesize that "language minority students placed in ESL may receive academic preparation different from that of their peers not placed in ESL and that ESL placement may actually preclude rather than ensure equity in curricular access." (2010, p.87). Callahan et al. found that students enrolled in ESL are 49% less likely to enroll in college preparatory science courses, 36% less likely to enroll in college preparatory math courses, and 56% less likely to enroll in college level social science courses (2010, p.96). They also found that students enrolled in ESL did worse than students not enrolled in ESL academically both on standardized tests, four points lower on a math test, and 0.18 less of a grade point for GPA (2010, p.101). The researchers also found that "long-term ELL language minority students – namely, those who are enrolled in ESL beyond an initial period when they are learning English – do not benefit from and may even be hindered by ESL placement" (Callahan et al., 2010, p. 104). Because of the potential negative findings in this study Callahan et al. encourage further study on the effect of ESL placement on ELLs' academic achievement.

Large scale studies. Researchers in both the English-only ESL and DL camps outline methodological issues in many of the studies describing the effectiveness of English language instructional programs. These methodological issues include: comparison or control group, unclear definitions of the treatment, small sample sizes, flawed conceptual frameworks, and researcher bias (Conger 2010; Ramirez et al., 1991;

Rolstad, Mahoney &Glass, 2008; Rossell 2003; Rossell & Baker, 1995; Slavin & Cheung, 2005). Several researchers have conducted large scale studies and meta-analyses to address these weaknesses in the literature. These larger studies provide insight into how difficult it is to effectively study instructional treatments for ELLs.

A study often cited by proponents of dual language programs is the one conducted by Ramirez et al. in 1991. The study was commissioned by the department of education and was one of the first longitudinal studies to analyze the effectiveness of English-only and dual language instructional programs for ELLs. Ramirez et al. (1991) defined structured English immersion programs as those where the formal language of instruction was English, native language was used on an informal limited basis only to clarify concepts, and subject matter instruction was taught in English. DL programs were defined as programs where teachers used both the native language (L1) and English (L2) for instruction and the use of each language is clearly defined, there was both L1 and L2 language arts component, math was taught in L1, and cultural sensitivity was reflected in teaching and instructional materials (Ramirez et al., 1991). Because of the large 2,352 N count in the study, geographic distribution, and large number of 554 classrooms, Ramirez et al. were confident about generalizing their findings on DL and structured English-only ESL programs (1991, p.80). However, in this large study, the researchers caution that the generalizability is for Spanish speaking ELLs because that was their sample (Ramirez et al., 1991). In kindergarten and first grade, students in English-only programs outperformed students in the DL programs, but by grade 3 the DL students were performing better than their peers in English-only programs. This finding is significant since students in the DL group were not receiving the same amount of English instruction

as their peers. This finding reinforces the theory that the more literate a student is in their first language the better they are able to learn a second (Cummings, 1992; Krashen, 2008; Lindholm-Leary, 2001). Ramirez concluded that a DL program

> does not interfere or delay [ELLs] acquisition of English language skills, but helps them to 'catch-up' to their English speaking peers in English language arts, English reading and math. In contrast, providing [ELLs] with almost exclusive instruction in English does not accelerate their acquisition of English language arts, reading or math, i.e., they do not appear to be 'catching up.' The data suggest that by grade six, students provided with English-only instruction may actually fall further behind their English-speaking peers" (Ramirez, 1992, p.1).

As important as the Ramirez et al. study was for DL language proponents, the work done by Rossell and Baker is cited just as often by the proponents of English-only ESL instruction. Rossell and Baker's study challenged the findings of Ramirez et al. and other researchers who had found positive results for DL programs, called into question second language acquisition principles, and cited the ideological bias of researchers studying ELL instructional programs.

Rossell and Baker (1996) conducted a narrative analysis of the research and found that ELLs achieved at higher levels in English-only ESL classrooms. Rossell and Baker selected studies for review that had both a random assignment group and a control group; had targeted assignments, a comparison group and no other educational treatments; or could statistically account for the other treatments the students were receiving. It is important to note that Rossell and Baker did not use studies showing the effectiveness of

dual language methods if the studies "merely described what went on in the bilingual education classroom rather than comparing students in bilingual education to similar students no in bilingual education" (1996, p.15).

The Ramirez et al. study concluded that late exit bilingual education, or DL programs are superior to structured English-only immersion programs. However, Rossell and Baker reviewed the evidence and derived a different conclusion, "Bilingual education may be superior to all English instruction in the very beginning when a student literally knows no English, but as the student's English language knowledge increase and English becomes more comprehensible, time-on-task in English becomes more important because it becomes *effective* time-on task" (1996, p.30). After identifying 72 studies that met their methodology criteria, Rossell and Baker found "no consistent research support for transitional bilingual education as a superior instructional practice improving the English language achievement of limited English proficient children" (1996, p.19). In terms of reading achievement Rossell and Baker found that 78% of the studies showed bilingual education no different or worse for ELLs than doing nothing at all (1996, p.20). In math it was 91% (1996, p.21). Structured immersion, where the ESL teachers teaches in English and sparingly uses some native language phrases to supplement understanding outperformed bilingual education. Bilingual education students did 83% worse and 17% showed no improvement in reading (1996, p. 21).

Rossell and Baker also question the validity of one of the major tenets of second language acquisition, the facilitation theory or threshold theory promoted by Cummings that states that the more developed the first language is the easier it is for a student to learn the second language. According to Rossell and Baker, there is no "underlying

psychological mechanism that accounts for the facilitation effect. Rather than being deduced from well established mental processes, the facilitation effect has to be accepted as a fundamental characteristic of the brain itself. At least parts of it, if not the whole hypothesis, are untestable" (1996, p.31). Rossell and Baker then challenge the evidence of Cummings and Collier, two prominent dual language proponents. Rossell and Baker state that the evidence Cummings provides for supporting his theory is "either trivial" or "just plain contrary" (1996, p. 29) and go on to cite major methodological issues with the research that Cummings cites to support his theory. Rossell and Baker state that Collier's research contradicted the facilitation theory because it did not show that students who had more command of their native language learned English at a faster rate than students whose first language was not as well developed. For Rossell and Baker, second language acquisition principles are not measurable and therefore cannot be used as the justification for an instructional model. Rossell and Baker do concede that bilingual education has some positive psychological effects and see the value in making the students feel welcomed and their second language valued so they are more motivated to come to school and learn. However, Rossell and Baker do not advocate the use of DL programs as a more effective way to teach ELLs.

Slavin and Cheung conducted a systematic review of the literature using a "best-evidence synthesis" where the researches attempt to "discover how much scientific basis there is for competing claims about the effects of bilingual as opposed to English-only programs" (2005, p. 248). Slavin and Cheung define English-only instruction, what they call English immersion, as instruction where "English language learners are expected to learn English from the beginning, and their native language plays little or no role in daily

reading lessons. Formal or informal support is likely to be given to ELLs to help them cope in an all-English classroom" (2005, p.250). They define DL programs as "a paired bilingual model, in that they [ELLs] learn to read both in English and in their native langue at different times each day" (2005, p.251).

Like Ramirez et al. and Rossell and Baker, Slavin and Cheung cite several problems with the previous research studies including comparison of programs at inappropriate times. One example of an inappropriate comparison is to compare a DL program where students do not begin getting the majority of their English instruction until fourth grade, but assess their progress and the progress of an English-only ESL group at first grade. Another issue is inappropriate assessment measures. An example of that may be using a pre-test in English for students that are in the first years in a bilingual program which could underestimate their skills in English while a Spanish pretest for English-only students would underestimate their skills in Spanish (Slavin & Cheung, 2005, p.251). Slavin and Cheung also cite selection bias for programs where students may be selected for one program or another by parent preference rather than the needs of the student. Schools engage in negative selection and that may result in placement of ELLs less capable to read in their native language to be placed in DL classes while more successful ELLs who can read in their own language are placed into English-only classes, thus stacking the deck against the DL students.

Slavin and Cheung also cite flaws in academic publishing and authoring of studies adding to the problem of objective reviews of ELL instructional methods. They cite the "'file drawer problem' the fact that studies showing no differences are less likely to be published or to otherwise come to light" (2005, p. 252). The researchers also cite that

bilingual education studies too often "say too little about the bilingual and immersion programs themselves or the degree or quality of implementation of the program. Yet bilingual models can vary substantially in quality, amount of exposure to English in and out of school, teachers' language facility, time during the school day, instructional strategies unrelated to language of instruction, and so on" (2005, p.253). Slavin and Cheung outline the ideal adequate study to compare the two treatments as having "(a) randomly assign a large number of children to be taught in English or their native language; (b) pretest them in their native language when they begin to be taught differentially, either in their native language or in English (typically in kindergarten); (c) follow them long enough for the latest-transitioning children in the bilingual condition to have completed their transition to English and have been taught long enough in English to make a fair comparison; and (d) collect data throughout the experiment to document the treatments received in all conditions. Unfortunately, only a few, very small studies of this kind have ever been carried out. As a result, the studies that compare bilingual and English-only approaches must be interpreted with great caution." (2005, p. 253).

Slavin and Cheung used the best-evidence synthesis and had five criteria: comparing students taught reading in bilingual classrooms and English immersion classes, students were assigned randomly or were given a pretest or other measure to establish comparability, subjects were all ELLs, dependent variables included quantitative measures of reading like standardized tests, and the treatment duration was at least one school year (2005, p.254-255). Slavin and Cheung found that few studies corroborated Rossell and Baker's work when reviewing the studies on how to best instruct ELLs. Slavin and Cheung evaluated Rossell and Baker's study using their best-

evidence synthesis design and found that Rossell and Baker did not select methodologically sound studies for their review. Slavin and Cheung had several important conclusions from their best-evidence synthesis: first is that there are too few high-quality studies and second, out of the 17 studies they found acceptable "12 found effects favoring bilingual education" and "none of the studies found results favoring English immersion" (2005, p.273).

Rolstad et al. conducted a meta-analysis to study effectiveness of instructional programs for ELLs and found that bilingual programs, such as dual language immersion, were consistently more effective than traditional English-only programs and had a positive effect of 0.23 standard deviations (2005, p.3). The researchers found that many of the narrative studies comparing the effectiveness of DL vs. ESL had inconsistent definitions of the programs and that led to results that could not be compared since programs with the same name had different definitions. For their study, Rolstad et al. focused on studies post 1985 and followed an approach of including as many studies as possible so the researchers could "probe more deeply into the distribution of study results to understand why some may find a stronger advantage for a particular program than another" (2005, p. 580). The researchers concluded that DL programs were more effective in increasing ELLs' academic achievement in both English and their native language.

While there are studies citing the effectiveness of both ESL and DL programs, a constant criticism is the lack of a strong conceptual framework for the instructional methods. Mitchell and Myles surmise it may be due to the fact that "…the findings for second language acquisition research are not sufficiently secure, clear and uncontested,

across broad enough domains, to provide straightforward prescriptive guidance for the teacher (nor, perhaps, will they ever be so). They are not generally presented and disseminated in ways accessible and meaningful to teachers…" (2004, p.261). The lack of consensus in second language acquisition research highlights one of the complexities in determining the effectiveness of English language programs, that there is no clear best practice for instructing ELLs. Researchers need to continue to study the second language acquisition process as well as the instructional programs using larger sample sizes and more strict experimental methods to determine what approach best serves ELLs.

CHAPTER 3

Methodology

This chapter provides the methodology applied in this study and is organized as follows: the purpose of the study, the research questions, the study design, the treatment, the treatment validity, the population and sample, the data collection, and the description of the data analysis procedures.

Purpose of the Study

The purpose of the study was to determine the extent to which pre-test proficiency level, gender, race, and identification as a student with disabilities (SWD) account for the variability in academic achievement for students enrolled in an ESL or a DL instructional program. The research questions examined the extent to which the independent variables accounted for variability in overall, reading, writing, listening, and speaking academic achievement for students enrolled in different instructional programs.

Research Questions

1. To what extent does pre-test proficiency level, gender, race, or SWD explain variability of overall academic achievement for students enrolled in an ESL program versus students enrolled in a DL program?
2. To what extent does pre-test proficiency level, gender, race, or SWD explain variability of reading academic achievement for students enrolled in an ESL program versus students enrolled in a DL program?

3. To what extent does pre-test proficiency level, gender, race, or SWD explain variability of writing academic achievement for students enrolled in an ESL program versus students enrolled in a DL program?
4. To what extent does pre-test proficiency level, gender, race, or SWD explain variability of listening academic achievement for students enrolled in an ESL program versus students enrolled in a DL program?
5. To what extent does pre-test proficiency level, gender, race, or SWD explain variability of speaking academic achievement for students enrolled in an ESL program versus students enrolled in a DL program?

Research Design

The study implemented a quasi-experimental design to determine the effects pre-test proficiency level, gender, race, and SWD have on the academic achievement of students enrolled in an ESL or a DL instructional program.

The independent selection variables were the instructional programs, ESL or DL. Pre-test proficiency level, gender, race, and SWD were the independent variables whose variability was analyzed through stepwise multiple linear regression analyses. The pre-test proficiency level was chosen since it is usually the most highly correlated variable to the post-test score (Trochim, 2005). Gender and race are variables not usually studied in this population and were added to add to the knowledge base. The reason race was chosen is that for NCLB, Latino/Hispanic is a subgroup counted for AYP, even though it is an ethnicity not a race. The students in the samples were all identified as Latino/Hispanic, but some students also identified themselves with several races, White and Black/African American. Since students that identify themselves both as

Latino/Hispanic and a minority race may be counted in multiple NCLB subgroups, it was important to analyze how race impacted academic outcomes. The last independent variable was identification as a student with disabilities (SWD). Little research has been focused on determining effective programs for ELLs with disabilities (Sáenz et al., 2005). Because there are significant numbers of ELLs who are also identified as students with disabilities, this variable was important to analyze.

The dependent variables were post-test proficiency levels for overall, reading, writing, listening, and speaking from the Assessing Comprehension and Communication in English State to State for English language learners (ACCESS for ELLs) assessment. The ACCESS is described in the sources of data section later in this chapter.

Treatment

The two treatments employed in the study were the ESL and DL instructional programs. These two instructional methods are extremely popular methods for teaching ELLs and studies show each program can increase academic achievement and English language proficiency.

In traditional ESL instruction the students learn English and academic content using only English as the language of instruction. Programs that focus on developing literacy using only English include programs in which students work only on English grammar, vocabulary, and communication skills and programs that blend both content and English instruction in an English only environment (Office of English Language Acquisition, 2008). To teach students academic content, traditional ESL programs may employ sheltered English instruction where the content is modified to make it more

comprehensible (Rossell, 2005; Short et al., 2011; Smith et al., 2008). Teachers using sheltered English instruction to teach content will have both content and language objectives. As students become more proficient in English, they transition out of ESL classes and move into classes that are mainstream.

In this study ESL instruction was provided in two main formats:

- small sheltered English classes where ELLs with lower proficiency levels receive instruction in English focused on grammar, vocabulary, and communication skills, and
- ESL inclusion where ESL teachers come into the mainstream classrooms and provide clarification, minimal translation, and other supports to ELLs so they can actively participate in a mainstream English-only class

(Office of English Language Acquisition, 2008; Violand-Sanchez, Tabatabai, Van Horne, & Varela, 2008)

The ESL program provided for students in this county was well established and implemented instructional techniques frequently cited (Calderón et al., 2011; Carlo et al., 2005; Collier, 1995; Echeverria et al., 2006) as effective for second language learners, such as:

- Integrated instruction in oral language, reading, and writing.
    - An example is a weeklong unit on the five senses where students practice pronouncing the five senses (touch, taste, smell, hear, see), participate in a

guided reading activity on the purpose of the five senses, and create a diagram of the five senses writing the words in the correct locations.

- Integrated language and content instruction with a thematic approach.
    - An example is a thematic unit on the thirteen colonies that consists of objectives based on the Virginia Standards of Learning as well as objectives that address language instruction. The unit includes language objectives that address spelling the thirteen colonies correctly and writing a sentence describing an important economic activity of each colony.
- Lessons and units that promote concept development.
    - This bullet is particularly important in content lessons. An example is a lesson on addition where students use manipulatives to learn the concept of addition before applying it to problems.
- Scaffolding of instruction to promote higher level of learning.
    - An example is building on the singular and plural forms of several nouns and having students complete sentences to demonstrate they understand the proper use of singular and plural forms.
- Providing ample opportunities for practicing language skills.
    - Teachers can provide opportunities for practicing language skills by checking for understanding through oral questioning. Students have a chance to practice speaking as well as to demonstrate understanding.

(Violand-Sanchez, et al., 2007, pp. 21-22)

In a DL classroom, ELLs receive content and language instruction in both their native language and English. DL programs implement the major tenets of second language acquisition theory which assumes that students who become highly proficient in their native language become successful second language learners. Providing a solid foundation in the native language through instruction and immersion in the native language is an integral part of DL instruction. DL programs have three major instructional components that implement second language acquisition principles into classroom instruction; including equal numbers of ELLs and English dominant students in the classroom, providing academic instruction to an integrated group of ELLs and English dominant students; and providing core academic content to both groups of students in both languages, and by third grade literacy instruction in both languages (Howard & Christian, 2002; Lindholm-Leary, 2001).

The DL program in this study applied a 50:50 instructional model for grades K-5 where students spent half the instructional day in English and half in Spanish throughout the length of the program (Christian et al., 2000; Lindholm-Leary, 2004). The DL program in this study had several program goals cited by researchers as necessary to develop effective bilingual education programs (Christian et al., 2000; Collier &Thomas, 2001; Lindholm-Leary &Block 2010). These program goals included:

- the development and maintenance of Spanish and English proficiency in listening, speaking, reading, and writing;

- the high academic achievement in content areas in both languages;

- cross cultural understanding;
- the integration of students with a wide range of abilities;
- and a positive environment in which the minority language (Spanish) is respected and celebrated

(Forbes-Ullrich, Perdomo, 2005, p.5)

The 50:50 DL program in this study was modeled after the research done by Lindholm-Leary (2001). The principle features of the program were as follows:

- Instructional design and features that identify the duration of the instructional program, specify the exposure to optimal dual-language input, and provide opportunities to use and promote language output.
    - An example of this is the dissemination of the instructional design and program outline to teachers, parents, and students via the county website and also in hardcopy.
- Focus on academic criteria.
    - An example of this focus is the criteria for the content classes that address the essential knowledge, skills, and understandings outlined in the Virginia Standards of Learning and assessed at the end of the year with high-stakes state mandated assessments.
- Integrated language arts instruction in both languages.

  - An example of this integration is the course sequence for the DL program. The DL program provided Spanish and English language arts classes starting in the third grade. However, explicit language instruction in both languages began at kindergarten.

- Separation of languages for instruction.
  - An example of separation of languages is found in the daily schedule. Aside from Spanish and English language arts, teachers in the DL program taught science and history in Spanish and math in English.

- Ratio of the use of English to the target language.
  - An example of applying a proper ratio can be found in the DL model used. The DL program employed a 50:50 model with half the day taught in English and half the day in taught Spanish.

- Literacy instruction in two languages.
  - An example of this can be found in the program outline. Dual language instruction began in kindergarten with more formalized English language arts and Spanish language arts in upper elementary grades.

- Diverse student body (class composition is racially-diverse and also includes students with special needs).

- An example of attempting to attain a diverse student body can be found in the matriculation system of the DL program. The DL program conducts an annual lottery where ELL and language dominant students are chosen in equal percentages for enrollment in the program. English dominant and Spanish dominant students were needed to be peer tutors and provide opportunities for practicing the second language.

- Home/school collaboration.
    - An example of this can be found in the expectations handed to parents when their students enroll in the program. This was an important piece in the program and parents were expected to participate in school activities as well as support their child at home.

(Forbes-Ullrich, Perdomo, 2005, p.5)

The DL program was in two elementary schools in the county while the ESL program was in 14 schools. The treatments were well documented and outlined for teachers, parents, and students. The teachers were supported through mandated training, volunteer staff development opportunities, and central office resources. The county in which these programs were housed had a large ELL population. Having well established programs in a county with enough resources to deliver to different instructional treatments for ELLs made it an ideal setting to conduct this study.

Treatment Validity

To ensure that the teachers were implementing DL and ESL instructional programs with fidelity, the researcher asked to conduct two observations in the DL and ESL schools. The request to conduct classroom observations was originally granted by the county, but then rescinded due to other internal program evaluations being conducted during the study period. The researcher then submitted questions to the Office of Planning and Evaluation to be answered by the directors of the ESL and DL programs. The Office of Planning and Evaluation provided confirmation that the teachers of the students selected for the study implemented the program with fidelity. However, the Office of Planning and Evaluation did not provide written answers to the researcher's questions. Without the unannounced observations the researcher cannot personally account for any variability due to the program implementation and can only report that the county Office of Planning and Evaluation stated that the programs were being implemented with fidelity.

Population and Sample

The population consisted of fourth and fifth grade ELL elementary students from a suburban county outside a major city in the Mid Atlantic who were enrolled in an instructional program to learn English. The county had two DL elementary schools, 14 elementary schools that implemented English-only ESL instructional programs, and five schools that provided first language support, a short-term transitional bilingual program.

The Office of Planning and Evaluation for the county identified the fourth and fifth grade ELLs at the two DL schools. All fourth and fifth grade ELLs at the two DL schools served as the sample for students receiving DL instruction. The Office of

Planning and Evaluation then identified fourth and fifth grade ELLs receiving ESL instruction at the other 14 schools. The Office of Planning and Evaluation provided the researcher with a study sample by selecting students who were receiving an ESL instructional program and were comparable to students in the DL programs on native language and ethnicity. The final sample consisted of all fourth and fifth grade ELLs enrolled in either an ESL or DL instructional program whose native language was Spanish and were Latino.

T-Tests were conducted to analyze whether the mean differences between the ESL and DL subgroups in the sample were statistically significant and to determine whether the two subgroups were comparable in the categories of pre-test proficiency level, race, gender, and SWD. The data from the T-Tests are shown in Table 1 for the continuous independent variables of pre-test level (overall, reading, writing, listening, speaking), and in Table 2 for the independent nominal variables of gender, race (Black/African American, White), and SWD.

Table1

Independent Samples T-Test Comparing Continuous Variables between ESL and DL Subgroups

		Levene's Test for Equality of Variances		T-Test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
Overall PL	Eq.var. assumed	4.189	.042	-1.443	244	.150	-.2359	.1635	-.5579	.0862
	Eq.var. not assumed			-1.567	173	.119	-.2359	.1505	-.5329	.0612
Reading PL	Eq. var. assumed	4.836	.029	-.926	245	.356	-.1740	.1880	-.5444	.1963
	Eq. var. not assumed			-1.009	174	.315	-.1740	.1726	-.5147	.1666
Writing PL	Eq.var. assumed	11.061	.001*	-1.803	244	.073	-.2413	.1338	-.5049	.0223
	Eq.var. not assumed			-1.953	172	.052	-.2413	.1235	-.4851	.0025
Listening PL	Eq. var. assumed	5.352	.022	-1.485	245	.139	-.2213	.1490	-.5148	.0722
	Eq.var. not assumed			-1.695	195	.092	-.2213	.1306	-.4789	.0363
Speaking PL	Eq.var. assumed	3.285	.071	-.463	245	.643	-.0971	.2095	-.5096	.3155
	Eq. var. not assumed			-.488	164	.626	-.0971	.1988	-.4896	.2954

Note. Eq. abbreviation for equal, var. abbreviation for variance, PL abbreviation for proficiency level

* p<.005 for Levene's Test for Equality of Variances

**p<.05 for T-Test for Equality of Means

As can be observed in Table 1, the pre-test proficiency level means for all five academic areas (overall, reading, writing, listening, speaking) were not significantly different between students in ESL or DL. The Levene's test for Equality of Variances needs an observed significance of less than .005 to prove the null hypothesis (Norusis, 2005). In Table 1, writing proficiency level was significant at .001. Since the Levene's test was significant, then the T-Test for the Equality of Means was used to validate that there was truly a significant difference between the means of the ESL and DL subgroups in writing. The T-Test writing proficiency level mean significance level was .052, greater than the .05 mean needed for the mean to be significantly different. Therefore, the ESL and DL subgroups were comparable in pre-test level means for all academic areas.

Table 2

Independent Samples T-Test Comparing Nominal Variables between ESL and DL Subgroups

		Levene's Test for Equality of Variances		T-Test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
Gender	Eq. var. assumed	.072	.789	.119	282	.905	.008	.067	-.123	.139
	Eq. var. not assumed			.119	138	.905	.008	.067	-.124	.140
Race - Black/African American	Eq. var. assumed	.208	.649	-.228	282	.820	-.003	.014	-.030	.024
	Eq. var. not assumed			-.214	123	.831	-.003	.015	-.032	.026
Race - White	Eq. var. assumed	2.130	.146	.702	282	.483	.041	.059	-.074	.156
	Eq. var. not assumed			.718	145	.474	.041	.057	-.072	.154
SWD	Eq. var. assumed	4.432	.036	-1.181	282	.239	-.074	.063	-.197	.049
	Eq. var. not assumed			-1.152	132	.251	-.074	.064	-.201	.053

Note. Eq. abbreviation for equal, var. abbreviation for variance, PL abbreviation for proficiency level, SWD abbreviation for students with disabilities
* p<.005 for Levene's Test for Equality of Variances
**p<.05 for T-Test for Equality of Means

As can be observed in Table 2, none of the nominal variable means were significantly different between ESL or DL students. The Levene's test for Equality of Variances needs an observed significance of less than .005 to prove the null hypothesis (Norusis, 2005). In Table 2, the gender significance was .789, the race (Black/African American) significance was .649, the race (White) significance was .146, and the SWD significance is .036. All the significance levels were greater than .005, so the means for the nominal independent variables were comparable in the DL and ESL subgroups.

Sources of Data

The study used archival achievement data provided by the county. The assessment data came from the Assessing Comprehension and Communication in English State to State for English language learners (ACCESS for ELLs). The World-Class Instructional Design and Assessment (WIDA) consortium developed the ACCESS by asking their consortia members to review and approve standards that focus on academic English and then taking these agreed upon standards and developing an assessment that measures listening, speaking, reading and writing (Bauman, Boals, Cranley, Gottlieb, & Kenyon, 2007). The test yields English language proficiency levels 1-6, with one corresponding to minimal proficiency and six corresponding to maximum proficiency. For students in the study sample, students with WIDA levels 1-4 were still considered ELLs while students with levels 5-6 were reclassified as proficient. Figure 2 describes the WIDA levels in terms of linguistic complexity, vocabulary usage, and language control.

Figure 2 WIDA Performance Definitions

WIDA PERFORMANCE DEFINITIONS

	Level 1 Entering	Level 2 Beginning	Level 3 Developing	Level 4 Expanding	Level 5 Bridging	Level 6 Reaching
Linguistic Complexity	Pictorial or graphic representation of the language of the content areas	General language related to the content areas	General and some specific language of the content areas	Specific and some technical language of the content areas	Specialized or technical language of the content areas	Specialized or technical language reflective of the content areas at grade level
Vocabulary Usage	Words, phrases or chunks of language when presented with one-step commands, directions, WH-, choice or yes/no questions, or statements with sensory, graphic or interactive support	Phrases or short sentences	Expanded sentence in oral interaction or written paragraphs	A variety of sentence lengths of varying linguistic complexity in oral discourse or multiple, related sentences or paragraphs	A variety of sentence lengths of varying linguistic complexity in extended oral or written discourse including stories, essays or reports	A variety of sentence lengths of varying linguistic complexity in extended oral or written discourse as required by the specified grade level
Language Control		Oral or written language with phonological, syntactic or semantic errors that often impede the meaning of the communication when presented with one to multiple-step commands, directions, questions or a series of statements with sensory, graphic or interactive support	Oral or written language with phonological, syntactic or semantic errors that may impede the communication , but retain much of its meaning, when presented with oral or written narrative or expository descriptions with sensory, graphic or interactive support	Oral or written language with minimal phonological, syntactic or semantic errors that do not impeded the overall meaning of the communication when presented with oral or written connected discourse with sensory, graphic or interactive support	Oral or written language approaching comparability to that of proficient English peers when presented with grade level material	Oral or written communication in English comparable to proficient English peers

Arlington Public Schools (2012) – ESOL/HILT WIDA Performance Definitions - http://apsva.us/Page/17260

The ACCESS differs from other large scale ELL assessments because it is based on WIDA standards, emphasizes academic English, has independent sub domains, is vertically aligned, and complies with NCLB mandates (Bauman et al., 2007). The test uses content standards in social studies, math, science, and English to assess students' English language proficiency.

This approach to testing English within other content domains was adopted by the county in this study in the 2009-2010 school year. The test was first used in the spring of 2010. The county instituted the use of the ACCESS to determine English language proficiency levels in the fall and spring in the 2010-2011 school year. The fall 2010 ACCESS scores served as the pre-test measure and the spring 2011 ACCESS scores served as the post-test measure. Before the ACCESS, English language proficiency level was assessed by using several measures including standardized tests, teacher observations, and student samples. The implementation of a new measure to assess English language proficiency may have attributed to the lack of correlation between instructional program and academic achievement gains, as well as the lack of correlation between the pre-test and the post-test.

The reason for using the ACCESS was that the test was designed for ELLs, so it should measure student scores more appropriately than a test designed for native English speakers. The pre-test was used as an independent variable and the post-test was used as the dependent variable to measure effects on achievement. This was done to strengthen the study design; however, there are several potential issues with using pre-tests and post-tests to analyze the effects of a treatment. One potential issue is that assessment is not sufficiently aligned to treatment. Studies attempting to measure the effects of

achievement with student populations often use state mandated tests. Using a standardized test versus an assessment closely aligned to the intervention may underestimate the effect of the treatment (Olsen, Unlu, Jaciw, & Price, 2011). There is also a potential effect if the assessment is used for high stakes decisions. Students may have a positive motivation to take the assessment if it has implications to their academic future and negative motivation if the assessment does not count for the students (Olsen et al., 2011). For this study, the standardized assessment was used for several purposes. The scores determined the English proficiency level of students and the potential reclassification as English proficient. The assessment was also used as a substitute for the state mandated reading assessments for students at the lower WIDA proficiency levels (1-2). The effects in selecting a standardized test as a study measure were considered when analyzing the results.

Another potential issue was the strength of the pre-test/post-test correlation because the sample was homogeneous. Ethnicity is an example of homogeneity for this sample; all students were Latino whose native language was Spanish. The pre-test/post-test correlation may have been weakened in this population due to the homogeneity of the sample. In a homogeneous group "small changes in student performance (such as changes due to random/measurement error) can result in large changes in relative position, or rank within the sample…Given that the pretest-posttest correlation coefficients are an indication of relative consistency of ranks in a sample, when a study focuses on a relatively homogenous subgroup, correlations will be attenuated." (Cole, Haimson, Perez-Johnson, & May, 2011, p. 12). The potentially weakened correlation was considered in the findings.

Data Analysis

Descriptive statistics including minimum, maximum, mean, standard deviation, and variance were provided for all the independent and dependent variables. A stepwise multiple linear regression analysis was utilized to determine to what extent pre-test proficiency level, gender, race, or SWD accounted for variability in overall academic achievement, reading academic achievement, writing academic achievement, listening academic achievement, and speaking academic achievement when selecting for students enrolled in an ESL program or students enrolled a DL program. The reason for using a stepwise multiple linear regression analysis was because there were multiple independent measures that the researcher wanted to test using the instructional programs as independent selection variables. Using a stepwise multiple regression analysis procedure provided a rank order of importance for the independent variables and provided a percentage of variability for each independent variable (Norusis, 2005). The stepwise multiple linear regression analysis provided the significant variables along with the combination of independent variables that contributed to the variability in achievement. These data provided information that was both statistically significant as well as practically significant when determining which instructional program to implement in a school division with limited resources.

CHAPTER 4

Presentation and Analysis of the Data

This chapter presents the data analysis of the effects that pre-test proficiency level, gender, race, and identification as student with disabilities (SWD) have on achievement for students enrolled in an ESL instructional program versus a DL instructional program. The chapter is organized as follows: descriptive statistics, correlation matrices and regression analyses, and summary of the data.

Descriptive Statistics for Independent and Dependent Variables

The following section reports the descriptive statistics for the independent and dependent variables used to analyze the effects on academic achievement for students enrolled in an ESL or a DL instructional program. The independent variables included pre-test proficiency levels (overall, reading, writing, listening, speaking), gender, race (Black/African American, White), SWD, and instructional program (ESL or DL). The dependent variables were the post-test proficiency levels (overall, reading, writing, listening, speaking). The nominal variables for ESL, DL, race, and SWD were coded as "1" for yes and as a "0" for no so they could be used like any other ordinal variable in the regression model (Norusis, 2005). Gender was coded "1" for female and "0" for male to examine the impact gender had on achievement. Since the pre-test and post-test proficiency level variables were continuous, there was no need to recode. Descriptive statistics including the minimum, maximum, mean, standard deviation, and variance for each of the independent and dependent variables are shown in Tables 3-5. The total N for

the sample was 283 students where 206 students belonged to the ESL sample and 78 students belonged to the DL sample.

Table 3

Descriptive Statistics for Total Sample Independent and Dependent Variables (Total N=283, ESL N=206, DL N=78)

Variable	N	Minimum	Maximum	M	SD	Variance
ESL	206	0	1	.73	.447	.200
DL	78	0	1	.27	.447	.200
Gender	283	0	1	.49	.501	.251
Race - Black/African American	283	0	1	.01	.102	.010
Race - White	283	0	1	.26	.440	.193
SWD	283	0	1	.33	.471	.222
Pre-test overall PL	245	1.1	6.0	4.40	1.183	1.400
Pre-test reading PL	247	1.0	6.0	4.67	1.358	1.845
Pre-test writing PL	246	1.0	6.0	3.73	.971	.942
Pre-test listening PL	246	1.0	6.0	4.96	1.080	1.166
Pre-test speaking PL	246	1.0	6.0	4.71	1.517	2.301
Post-test overall PL	245	1.7	6.0	4.59	.900	.810
Post-test reading PL	247	1.9	6.0	4.83	1.048	1.098
Post-test writing PL	246	1.0	6.0	4.15	.611	.374
Post-test listening PL	246	2.3	6.0	5.11	.780	.608
Post-test speaking PL	246	1.0	6.0	4.40	1.425	2.031

Note. PL abbreviation for proficiency level

As can be observed in Table 3, the ESL and DL standard deviations were the same at .447. The ESL mean was .73 while the DL mean was .27. The minimum and maximum pre-test and post-test proficiency levels ranged from 1.0 to 6.0. The mean for pre-test overall proficiency level was 4.40 with a standard deviation of 1.183 while the post-test overall proficiency level was 4.59 with a standard deviation of .900. The mean gain for overall proficiency was .19. For reading the pre-test proficiency level mean was 4.67 with a standard deviation of 1.358 while the post-test proficiency level mean 4.83 with a standard deviation of 1.048. The mean gain for reading proficiency was .16. For writing the pre-test proficiency level mean was 3.73 with a standard deviation of .971 while the post-test proficiency level mean was 4.15 with a standard deviation of .611. The mean gain for writing proficiency was .42—the largest gain for any of the proficiency levels studied. For listening the pre-test proficiency level mean was 4.96 with a standard deviation of 1.080 while the post-test proficiency level mean was 5.11 with a standard deviation of .780. The mean gain for listening proficiency was .15. For speaking the pre-test proficiency level mean was 4.71 with a standard deviation of 1.517 while the post-test proficiency level mean was 4.40 with a standard deviation of 1.425. The mean drop in proficiency was .31. This was the only proficiency level studied where there was a drop in proficiency.

Tables 4 and 5 show the descriptive statistics for independent and dependent variables broken down by instructional program. The data in these tables provide a more detailed look at two groups that composed the study sample.

Table 4

Descriptive Statistics for ESL Independent and Dependent Variables (Total N=206)

Variable	N	Minimum	Maximum	M	SD	Variance
Gender	206	0	1	.50	.501	.251
Race - Black/African American	206	0	1	.01	.098	.010
Race - White	206	0	1	.27	.446	.199
SWD	206	0	1	.31	.464	.215
Pre-test overall PL	170	1.1	6.0	4.32	1.248	1.559
Pre-test reading PL	172	1.0	6.0	4.61	1.438	2.069
Pre-test writing PL	171	1.0	5.6	3.65	1.020	1.040
Pre-test listening PL	171	1.0	6.0	4.89	1.169	1.366
Pre-test speaking PL	170	1.0	6.0	4.68	1.578	2.490
Post-test overall PL	170	1.7	6.0	4.52	.929	.863
Post-test reading PL	172	1.9	6.0	4.76	1.082	1.171
Post-test writing PL	171	1.0	6.0	4.10	.634	.402
Post-test listening PL	171	2.3	6.0	5.07	.789	.622
Post-test speaking PL	170	1.0	6.0	4.36	1.482	2.197

Note. PL abbreviation for proficiency level

As can be observed in Table 4, the ESL subgroup nominal variable means were comparable to the total sample nominal variable means in Table 3. The mean for gender was .01 higher than the total sample mean, the same as the total sample mean for race (Black/African American), .01 higher for race (White), and .02 lower for SWD. The pre-test and post-test proficiency level means for the ESL subgroup were lower than the total

sample proficiency level means. The greatest difference in pre-test proficiency level means were in overall and writing proficiency levels where the ESL subgroup mean was .08 less than the total sample proficiency level means. The least difference in pre-test proficiency level mean was in speaking proficiency level where the mean is .03 less than the total sample proficiency level mean. The greatest difference in post-test proficiency level means were in overall and reading proficiency levels where the ESL subgroup mean was .07 less than the total sample proficiency level means. The least difference in post-test proficiency level means were in listening and speaking where the ESL subgroup mean was .04 less than the total sample proficiency level mean.

Table 5

Descriptive Statistics for DL Independent and Dependent Variables (Total N=78)

Variable	N	Minimum	Maximum	M	SD	Variance
Gender	78	0	1	.49	.503	.253
Race - Black/African American	78	0	1	.01	.113	.013
Race - White	78	0	1	.23	.424	.180
SWD	78	0	1	.38	.490	.240
Pre-test overall PL	75	1.8	6.0	4.60	1.008	1.015
Pre-test reading PL	75	1.6	6.0	4.79	1.154	1.332
Pre-test writing PL	75	1.0	6.0	3.89	.827	.688
Pre-test listening PL	75	2.5	6.0	5.11	.827	.684
Pre-test speaking PL	76	1.9	6.0	4.78	1.377	1.896
Post-test overall PL	75	2.8	6.0	4.77	.797	.635
Post-test reading PL	75	2.4	6.0	5.00	.937	.877
Post-test writing PL	75	2.8	6.0	4.28	.527	.278
Post-test listening PL	75	3.0	6.0	5.21	.752	.565
Post-test speaking PL	76	1.8	6.0	4.56	1.258	1.582

Note. PL abbreviation for proficiency level

As can be observed in Table 5, the nominal variable means for the DL subgroup were comparable to the total sample nominal variable means in Table 3. The mean for gender and for race (Black/African American) was the same as the total sample mean, .03 lower for race (White), .05 higher for SWD. The pre-test and post-test means for the DL subgroup were higher than the total sample means. The greatest difference in pre-test

proficiency level mean was in overall proficiency where the DL mean was .20 higher than the total sample proficiency level mean. The least difference in pre-test proficiency level mean was in speaking where the DL mean is .07 higher than the total sample proficiency level mean. The greatest difference in post-test proficiency level mean was in overall proficiency level where the DL mean was .18 higher than the total sample proficiency level mean. The least difference in post-test proficiency level mean was in listening where the DL mean was .10 higher than the total sample proficiency level mean.

Analyses to Answer Research Questions

The following section reports data related to the research questions outlined in Chapter 3. Since all the independent variables means were comparable for students in both the ESL and DL subgroups (as stated in the Population and Sample section of Chapter 3, Tables 1 and 2), instructional program was used as the independent selection variable in the stepwise multiple linear regression analyses to analyze how the other independent variables attributed to the variance in academic achievement in overall, reading, writing, listening, and speaking for students enrolled in an ESL instructional program versus students enrolled in a DL instructional program. Correlations were included to provide information on the relationships among the independent variables. Stepwise multiple linear regression analyses were conducted for overall, reading, writing, listening, and speaking academic proficiency levels to determine the effects that pre-test proficiency level, gender, race, and SWD had on academic achievement for students enrolled in an ESL instructional program versus a DL instructional program. The alpha level for statistical significance was set at .05. The FIN was set at .50 and the FOUT was

set at 1.0 to determine the portion of variance explained by each independent variable and to analyze if there were indications of practical significance.

Question 1: To what extent does pre-test proficiency level, gender, race, or SWD explain variability of overall academic achievement for students enrolled in an ESL program versus students enrolled in a DL program? To analyze the relationships among the independent variables the researcher produced a correlation matrix using the Pearson correlation and conducted a stepwise multiple regression to answer question 1. Data from the matrix are presented in Table 6 for ESL and Table 7 for DL and data from the regression analyses are presented in Table 8 for ESL and Table 9 for DL.

Table 6

Correlation Matrix for Overall Academic Proficiency Level Using Pearson Correlation (ESL)

Variables	Post-test Overall Proficiency Level	Gender	Race - Black/African American	Race - White	Pre-test Overall Proficiency Level	SWD
Post-test Overall Proficiency Level	1.000	-.032	-.076	.035	.142	-.501
Gender	-.032	1.000	-.108	.007	-.098	-.113
Race - Black/African American	-.076	-.108	1.000	.068	-.067	.031
Race - White	.035	.007	.068	1.000	.032	.041
Pre-test Overall Proficiency Level	.142	-.098	-.067	.032	1.000	-.122
SWD	-.501	-.113	.031	.041	-.122	1.000

As can be observed in Table 6, there was a strong negative correlation between SWD and post-test proficiency level of -.501 for students enrolled in an ESL program.

The next strongest correlation was between pre-test proficiency level and post-test proficiency level at .142. The weakest correlations to post-test overall proficiency level were gender at -.032, race (Black/African American) at -.076, and race (White) at .035.

Table 7

Correlation Matrix for Overall Academic Proficiency Level Using Pearson Correlation (DL)

Variables	Post-test Overall Proficiency Level	Gender	Race - Black/African American	Race - White	Pre-test Overall Proficiency Level	SWD
Post-test Overall Proficiency Level	1.000	-.006	.002	-.153	-.084	-.598
Gender	-.006	1.000	-.112	.054	.042	-.185
Race - Black/African American	.002	-.112	1.000	-.063	-.018	.142
Race - White	-.153	.054	-.063	1.000	.080	.013
Pre-test Overall Proficiency Level	-.084	.042	-.018	.080	1.000	.072
SWD	-.598	-.185	.142	.013	.072	1.000

As can be observed in Table 7, there was a strong negative correlation between SWD and post-test proficiency level of -.598 for students enrolled in a DL program. The next strongest correlation was between race (White) and post-test proficiency level at -.153. Pre-test and post-test proficiency levels were correlated at a -.084. The weakest correlations to post-test overall proficiency level were gender at -.006 and race (Black/African American) at .002.

Table 8

Stepwise Multiple Regression Analysis of Effects of Pre-test Proficiency Level, Gender, Race, and SWD on Overall Achievement for Students Enrolled in an ESL Program (N=170)

Variable	r	r^2	Cum R	Cum R^2	R^2 Change	F Value	F Change	Sig. F Change
SWD	-.501	.251	.501	.251	.251	56.196	56.196	.000**
Pre-test Overall Proficiency Level	.006	.000036	.507	.257	.006	28.986	1.492	.224
Race (White)	-.003	.000009	.510	.260	.003	19.476	.684	.409
Race (Black/African American)	.003	.000009	.513	.263	.003	14.755	.700	.404
Gender	.008	.00064	.521	.271	.008	12.213	1.767	.186

** p<.0001
*p<.05

As observed in Table 8, SWD was the only significant variable accounting for 25.1% of the variability in overall academic achievement. The other nominal variables accounted for 1.4% of the variability in achievement; gender accounted for .8%, race (White) accounted for .3%, and race (Black/African American) accounted for .3%. The pre-test proficiency level accounted for .6% of the variability in achievement. The independent variables accounted for 27.1% of the variability in overall academic achievement for students enrolled in an ESL program.

Table 9

Stepwise Multiple Regression Analysis of Effects of Pre-test Proficiency Level, Gender, Race, and SWD on Overall Achievement for Students Enrolled in an DL Program (N=75)

Variable	r	r^2	Cum R	Cum R^2	R^2 Change	F Value	F Change	Sig. F Change
SWD	-.598	.358	.598	.358	.358	40.651	40.651	.000**
Race (White)	-.018	.000324	.616	.376	.018	21.962	2.461	.121
Pre-test Overall Proficiency Level	-.001	.000001	.617	.377	.001	14.491	.009	.754
Gender	-.009	.000081	.626	.386	.009	11.267	1.369	.246
Race (Black/African American)	.004	.000016	.630	.390	.004	9.065	.549	.461

** p<.0001
*p<.05

As observed in Table 9, SWD was the only significant variable accounting for 35.8% of the variability. The other nominal variables accounted for 3.1% of the variability in overall academic achievement; race (White) accounted for 1.8%, gender accounted for .9%, and race (Black/African American) accounted for .4%. For students enrolled in a DL program, the pre-test accounted for .1% of the variability in overall academic achievement. All of the independent variables accounted for 39% of the variability in overall academic achievement for students enrolled in a DL program.

Question 2: To what extent does pre-test proficiency level, gender, race, or SWD explain variability of overall reading achievement for students enrolled in an ESL program versus students enrolled in a DL program? To analyze the relationships among the

independent variables the researcher produced a correlation matrix using the Pearson correlation and conducted a stepwise multiple regression to answer question 2. Data from the matrix are presented in Table 10 for ESL and Table 11 for DL and data from the regression analyses are presented in Table 12 for ESL and Table 13 for DL.

Table 10

Correlation Matrix for Reading Academic Proficiency Level Using Pearson Correlation (ESL)

Variables	Post-test Reading Proficiency Level	SWD	Race - Black/African American	Pre-test Reading Proficiency Level	Race - White	Gender
Post-test Reading Proficiency Level	1.000	-.500	-.146	.121	.046	.003
SWD	-.500	1.000	.030	-.101	.056	-.100
Race - Black/African American	-.146	.030	1.000	-.054	.067	-.107
Pre-test Reading Proficiency Level	.121	-.101	-.054	1.000	.016	-.104
Race - White	.046	.056	.067	.016	1.000	.020
Gender	.003	-.100	-.107	-.104	.020	1.000

As can be observed in Table 10, there was a strong negative correlation between SWD and post-test proficiency level of -.500 for students enrolled in an ESL program. The next strongest correlation was between race (Black/African American) and post-test proficiency level at -.146. Pre-test and post-test proficiency levels were correlated at a .121. The weakest correlations to post-test overall proficiency level were race (White) at .046 and gender at .003.

Table 11

Correlation Matrix for Reading Academic Proficiency Level Using Pearson Correlation (DL)

Variables	Post-test Reading Proficiency Level	SWD	Race - White	Pre-test Reading Proficiency Level	Race - Black/African American	Gender
Post-test Reading Proficiency Level	1.000	-.523	-.170	-.107	.024	-.021
SWD	-.523	1.000	.013	.054	.142	-.185
Race - White	-.170	.013	1.000	.145	-.063	.054
Pre-test Reading Proficiency Level	-.107	.054	.145	1.000	-.090	.031
Race - Black/African American	.024	.142	-.063	-.090	1.000	-.112
Gender	-.021	-.185	.054	.031	-.112	1.000

As can be observed in Table 11, there was a strong negative correlation between SWD and post-test proficiency level of -.523 for students enrolled in a DL program. The next strongest correlation was between race (White) and post-test proficiency level of -.170. The pre-test and post-test proficiency levels had a correlation of -.107. The least correlated variables to the post-test were race (Black/African American) at .024 and gender at -.021.

Table 12

Stepwise Multiple Regression Analysis of Effects of Pre-test Proficiency Level, Gender, Race, and SWD on Reading Achievement for Students Enrolled in an ESL Program (N=172)

Variable	r	r^2	Cum R	Cum R^2	R^2 Change	F Value	F Change	Sig. F Change
SWD	-.500	.250	.500	.250	.250	56.625	56.625	.000**
Race (Black/African American)	-.017	.000289	.517	.267	.017	30.788	3.963	.048*
Pre-test Reading Proficiency Level	.004	.000016	.521	.271	.004	20.833	.945	.332
Race (White)	.006	.000036	.527	.277	.006	16.050	1.509	.221
Gender	-.003	.000009	.530	.280	.003	12.968	.741	.391

** p<.0001
*p<.05

As can be observed in Table 12, for students enrolled in an ESL program SWD was a significant variable accounting for 25% of the variability in reading achievement. Race (Black/African American) was also a significant variable accounting for 1.7% of the variability in reading academic achievement. The remaining two nominal variables accounted for .9% of the variability in reading achievement; race (White) accounted for .6%, and gender accounted for .3%. The pre-test accounted for .4% of the variability in achievement. All the independent variables combined accounted for 28% of the variability in reading achievement for students in an ESL program.

Table 13

Stepwise Multiple Regression Analysis of Effects of Pre-test Proficiency Level, Gender, Race, and SWD on Reading Achievement for Students Enrolled in an DL Program (N=75)

Variable	r	r^2	Cum R	Cum R^2	R^2 Change	F Value	F Change	Sig. F Change
SWD	-.523	.273	.523	.273	.273	27.416	27.416	.000**
Race (White)	-.024	.000576	.547	.297	.024	15.406	2.741	.102
Pre-test Reading Proficiency Level	-.003	.000009	.550	.300	.003	10.282	.323	.571
Race (Black/African American)	.007	.000049	.557	.307	.007	7.860	.719	.400
Gender	-.009	.000081	.566	.316	.009	6.506	1.062	.306

** p<.0001
*p<.05

As can be observed in Table 13, SWD was a significant variable accounting for 27.3% of the variability in reading achievement for students enrolled in a DL program. The remaining nominal variables accounted for 4% of the variability in achievement; race (White) accounted for 2.4%, race (Black/African American) accounted for .7%, and gender accounted for .9% . Pre-test proficiency level accounted for .3% of the variability in reading achievement. All independent variables combined accounted for 31.6% of the variability in reading academic achievement for students enrolled in a DL program.

Question 3: To what extent does pre-test proficiency level, gender, race, or SWD explain variability of overall writing achievement for students enrolled in an ESL program versus

students enrolled in a DL program? To analyze the relationships among the independent variables the researcher produced a correlation matrix using the Pearson correlation and conducted a stepwise multiple regression to answer question 3. Data from the matrix are presented in Table 14 for ESL and Table 15 for DL and data from the regression analyses are presented in Table 16 for ESL and Table 17 for DL.

Table 14

Correlation Matrix for Writing Academic Proficiency Level Using Pearson Correlation (ESL)

Variables	Post-test Writing Proficiency Level	SWD	Race - White	Race - Black/African American	Pre-test Writing Proficiency Level	Gender
Post-test Writing Proficiency Level	1.000	-.483	.075	-.067	.061	.050
SWD	-.483	1.000	.054	.030	-.024	-.105
Race - White	.075	.054	1.000	.066	.071	.017
Race - Black/African American	-.067	.030	.066	1.000	-.054	-.108
Pre-test Writing Proficiency Level	.061	-.024	.071	-.054	1.000	-.130
Gender	.050	-.105	.017	-.108	-.130	1.000

As can be observed in Table 14, there was a strong negative correlation between post-test proficiency level and SWD of -.483 for students in an ESL program. The rest of the nominal variables correlated to the post-test proficiency level as follows; race (White) at .075, race (Black/African American) at -.067, and gender at .050. The pre-test and post-test had a correlation of .061.

Table 15

Correlation Matrix for Writing Academic Proficiency Level Using Pearson Correlation (DL)

Variables	Post-test Writing Proficiency Level	SWD	Pre-test Writing Proficiency Level	Gender	Race - White	Race - Black/African American
Post-test Writing Proficiency Level	1.000	-.486	-.143	.120	-.082	.024
IEP Status	-.486	1.000	.112	-.185	.013	.142
Pre-test Writing Proficiency Level	-.143	.112	1.000	.005	.124	.086
Gender	.120	-.185	.005	1.000	.054	-.112
Race - White	-.082	.013	.124	.054	1.000	-.063
Race - Black/African American	.024	.142	.086	-.112	-.063	1.000

As can be observed in Table 15, there was a strong negative correlation between SWD and the post-test writing proficiency level of -.486 for students enrolled in a DL program. The pre-test was the next highest correlated variable at -.143. The rest of the nominal variables correlated as follows; gender at .120, race (White) at -.082, and race (Black/African American) at .024.

Table 16

Stepwise Multiple Regression Analysis of Effects of Pre-test Proficiency Level, Gender, Race, and SWD on Writing Achievement for Students Enrolled in an ESL Program (N=171)

Variable	r	r^2	Cum R	Cum R^2	R^2 Change	F Value	F Change	Sig. F Change
SWD	-.483	.233	.483	.233	.233	51.373	51.373	.000**
Race (White)	.010	.001	.493	.243	.010	27.026	2.287	.132
Race (Black/African American)	-.004	.000016	.497	.247	.004	18.253	.779	.379
Pre-test Writing Proficiency Level	.001	.000001	.498	.248	.001	13.719	.336	.563
Gender	-.001	.000007	.499	.249	.001	10.910	.003	.954

** p<.0001
*p<.05

As can be observed in Table 16, SWD was a significant variable accounting for 23.3% of the variability in writing achievement for students enrolled in an ESL program. The rest of the nominal variables accounted for a total of 1.5%; race (White) accounted for 1%, race (Black/African American) accounted for .4%, and gender accounted for .1%. The pre-test accounted for .1% of the variability in writing achievement. All the independent variables accounted for 24.9% of the variability in writing achievement for students enrolled in an ESL program.

Table 17

Stepwise Multiple Regression Analysis of Effects of Pre-test Proficiency Level, Gender, Race, and SWD on Writing Achievement for Students Enrolled in a DL Program (N=75)

Variable	r	r^2	Cum R	Cum R^2	R^2 Change	F Value	F Change	Sig. F Change
SWD	-.486	.236	.486	.236	.236	22.538	22.538	.000**
Pre-test Writing Proficiency Level	-.008	.000064	.494	.244	.008	11.608	.754	.388
Gender	.001	.000001	.495	.245	.001	7.677	.103	.749
Race (White)	-.004	.000016	.499	.249	.004	5.815	.419	.520
Race (Black/African American)	.010	.0001	.509	.259	.010	4.830	.918	.341

** p<.0001
*p<.05

As can be observed in Table 17, SWD was a significant variable that accounted for 23.6% of the variability in writing achievement for students in a DL program. The rest of the nominal variables accounted for 1.5% of the variability; race (Black/African American) 1%, Race (White) .4%, and gender .1%. The pre-test accounted for .8% of the variability in writing achievement. The independent variables accounted for 25.9% of the variability in writing academic achievement for students enrolled in a DL program.

Question 4: To what extent does pre-test proficiency level, gender, race, or SWD explain variability of overall listening achievement for students enrolled in an ESL program versus students enrolled in a DL program? To analyze the relationships among the independent variables the researcher produced a correlation matrix using the Pearson

correlation and conducted a stepwise multiple regression to answer question 4. Data from the matrix are presented in Table 18 for ESL and Table 19 for DL and data from the regression analyses are presented in Table 20 for ESL and Table 21 for DL.

Table 18

Correlation Matrix for Listening Academic Proficiency Level Using Pearson Correlation (ESL)

Variables	Post-test Listening Proficiency Level	SWD	Pre-test Listening Proficiency Level	Gender	Race - Black/African American	Race - White
Post-test Listening Proficiency Level	1.000	-.432	.168	-.046	-.030	.015
SWD	-.432	1.000	-.158	-.108	.031	.043
Pre-test Listening Proficiency Level	.168	-.158	1.000	-.118	-.087	.050
Gender	-.046	-.108	-.118	1.000	-.107	.010
Race - Black/African American	-.030	.031	-.087	-.107	1.000	.068
Race - White	.015	.043	.050	.010	.068	1.000

As can be observed in Table 18, SWD was highly correlated to post-test proficiency level at -.432 for students enrolled in an ESL program. The next highest correlated variables were pre-test and post-test at .168. The rest of the nominal variables correlated to the post-test as follows; gender at -.046, race (Black/African American) at -.030, and race (White) at .015.

Table 19

Correlation Matrix for Listening Academic Proficiency Level Using Pearson Correlation (DL)

Variables	Post-test Listening Proficiency Level	SWD	Pre-test Listening Proficiency Level	Gender	Race - White	Race - Black/African American
Post-test Listening Proficiency Level	1.000	-.505	.091	-.069	-.064	.028
SWD	-.505	1.000	-.098	-.185	.013	.142
Pre-test Listening Proficiency Level	.091	-.098	1.000	.119	-.054	.126
Gender	-.069	-.185	.119	1.000	.054	-.112
Race - White	-.064	.013	-.054	.054	1.000	-.063
Race - Black/African American	.028	.142	.126	-.112	-.063	1.000

As observed in Table 19, SWD had a strong negative correlation to post-test proficiency level for students enrolled in a DL program. Pre-test and post-test correlated at .091. The last three nominal variables correlated as follows; gender at -.069, race (White) at -.064, and race (Black/African American) at .028.

Table 20

Stepwise Multiple Regression Analysis of Effects of Pre-test Proficiency Level, Gender, Race, and SWD on Listening Achievement for Students Enrolled in an ESL Program (N=171)

Variable	r	r^2	Cum R	Cum R^2	R^2 Change	F Value	F Change	Sig. F Change
SWD	-.432	.187	.432	.187	.187	38.786	38.796	.000**
Pre-test Listening Proficiency Level	.012	.000144	.444	.199	.012	20.592	2.137	.146
Gender	-.007	.000049	.451	.206	.007	14.211	1.361	.245
Race (Black/African American)	-.001	.000001	.452	.207	.001	10.615	.066	.798
Race (White)	.001	.000001	.453	.208	.001	8.492	.203	.653

** p<.0001
*p<.05

As observed in Table 20, SWD was a significant variable accounting for 18.7% of the variability in listening academic achievement. The rest of the nominal variables accounted for .9% of the variability in listening achievement; gender .7%, race (Black/African American) .1% and race (White) .1%. The pre-test proficiency level accounted for 1.2% of the variability in achievement. All of the independent variables accounted for 20.8% of the variability in listening academic achievement for students enrolled in an ESL program.

Table 21

Stepwise Multiple Regression Analysis of Effects of Pre-test Proficiency Level, Gender, Race, and SWD on Listening Achievement for Students Enrolled in a DL Program (N=75)

Variable	r	r^2	Cum R	Cum R^2	R^2 Change	F Value	F Change	Sig. F Change
SWD	-.505	.255	.505	.255	.255	25.013	25.013	.000**
Pre-test Listening Proficiency Level	.002	.00004	.507	.257	.002	12.446	.164	.687
Gender	-.028	.000784	.535	.285	.028	9.488	2.911	.092
Race (White)	-.002	.00004	.537	.287	.002	7.084	.194	.661
Race (Black/African American)	.005	.000025	.542	.292	.002	5.746	.569	.453

** p<.0001
*p<.05

As can be observed in Table 21, SWD was a significant variable accounting from 25.5% of the variability in listening academic achievement. The rest of the nominal variables accounted for 3.5% of the variability in achievement; gender 2.8%, race (White) .2%, and race (Black/African American) .5%. The pre-test proficiency level accounted for .2% of the variability in achievement. All the independent variables accounted for 29.2% of listening variability for students enrolled in a DL program.

Question 5: To what extent does pre-test proficiency level, gender, race, or SWD explain variability of overall speaking achievement for students enrolled in an ESL program versus students enrolled in a DL program? To analyze the relationships among the independent variables the researcher produced a correlation matrix using the Pearson

correlation and conducted a stepwise multiple regression to answer question 5. Data from the matrix are presented in Table 22 for ESL and Table 23 for DL and data from the regression analyses are presented in Table 24 for ESL and Table 25 for DL.

Table 22

Correlation Matrix for Speaking Academic Proficiency Level Using Pearson Correlation (ESL)

Variables	Post-test Speaking Proficiency Level	SWD	Pre-test Speaking Proficiency Level	Race - White	Gender	Race - Black/African American
Post-test Speaking Proficiency Level	1.000	-.161	-.069	.050	-.048	.031
SWD	-.161	1.000	-.094	.041	-.113	.031
Pre-test Speaking Proficiency Level	-.069	-.094	1.000	-.033	-.013	-.151
Race - White	.050	.041	-.033	1.000	.007	.068
Gender	-.048	-.113	-.013	.007	1.000	-.108
Race - Black/African American	.031	.031	-.151	.068	-.108	1.000

As can be observed in Table 22, there was a negative correlation between SWD and post-test of -.161. The rest of the nominal variables correlated to the post-test as follows; race (White) at .050, gender at -.048, and race (Black/African American) at .031. The pre-test correlated to the post-test at -.069.

Table 23

Correlation Matrix for Speaking Academic Proficiency Level Using Pearson Correlation (DL)

Variables	Post-test Speaking Proficiency Level	SWD	Race - White	Gender	Pre-test Speaking Proficiency Level	Race - Black/African American
Post-test Speaking Proficiency Level	1.000	-.256	-.092	-.081	-.016	.000
SWD	-.256	1.000	.019	-.173	.054	.143
Race - White	-.092	.019	1.000	.060	.043	-.062
Gender	-.081	-.173	.060	1.000	-.107	-.110
Pre-test Speaking Proficiency Level	-.016	.054	.043	-.107	1.000	-.176
Race - Black/African American	.000	.143	-.062	-.110	-.176	1.000

As can be observed in Table 23, the greatest correlation between variables was SWD and post-test proficiency level at -.256. The rest of the nominal variables correlated with the post-test as follows; race (White) at -.092, gender at -.081, while race (Black/African American) did not correlate at all with the post-test proficiency level. The pre-test had a weak correlation at -.016.

Table 24

Stepwise Multiple Regression Analysis of Effects of Pre-test Proficiency Level, Gender, Race, and SWD on Speaking Achievement for Students Enrolled in an ESL Program (N=170)

Variable	r	r^2	Cum R	Cum R^2	R^2 Change	F Value	F Change	Sig. F Change
SWD	-.161	.026	.161	.026	.026	4.450	4.450	.036*
Pre-test Speaking Proficiency Level	-.020	.0004	.181	.046	.020	2.840	1.224	.270
Race (White)	.008	.000064	.189	.054	.008	2.055	.502	.479
Gender	-.002	.000004	.191	.056	.002	1.744	.818	.367
Race (Black/African American)	.001	.000001	.192	.057	.001	1.392	.027	.869

** p<.0001
*p<.05

As can be observed in Table 24, SWD was a significant variable accounting for 2.6% of the variability in speaking academic achievement. The rest of the nominal variables accounted for 1.1% of the variability; race (White) .8%, gender .2%, and race (Black/African American) .1%. The pre-test accounted for 2% of the variability in speaking achievement. All the independent variables accounted for 5.7% of the variability in speaking academic achievement for students enrolled in an ESL program.

Table 25

Stepwise Multiple Regression Analysis of Effects of Pre-test Proficiency Level, Gender, Race, and SWD on Speaking Achievement for Students Enrolled in a DL Program (N=76)

Variable	r	r^2	Cum R	Cum R^2	R^2 Change	F Value	F Change	Sig. F Change
SWD	-.256	.066	.256	.066	.066	5.189	5.189	.026*
Race (White)	-.004	.000016	.230	.070	.004	2.877	.595	.443
Gender	-.026	.000676	.256	.096	.026	2.312	1.168	.283
Pre-test Speaking Proficiency Level	-.001	.000001	.257	.097	.001	1.713	.009	.923
Race (Black/African American)	.001	.000001	.258	.098	.001	1.357	.028	.868

** p<.0001
*p<.05

As can be observed in Table 25, SWD was a significant variable accounting for 6.6% of the variability in speaking achievement. The rest of the nominal variables accounted for 3.1% of the variability in achievement; gender 2.6%, race (White) .4%, and race (Black/African American .1%. The pre-test accounted for .1% of the variability in speaking achievement. All the independent variables accounted for 9.8% of the variability in listening academic achievement for students in a DL program.

Further Analysis

The following section presents data for the percentage of variability in academic achievement accounted for by pre-test proficiency level, gender, race, and SWD, the

mean pre-test proficiency levels, the mean post-test proficiency levels, and the mean gain scores for students enrolled in an ESL instructional program and students enrolled in a DL program.

Table 26

Total Variances for Academic Areas, Mean Scores, and Mean Gain Scores for ESL and DL Programs

Academic Area	Percentage of Variability	Mean Pre-test Proficiency Level	Mean Post-Test Proficiency Level	Mean Gain Score
Overall ESL	27.1	4.32	4.52	.20
Overall DL	39.0	4.60	4.77	.17
Reading ESL	28.1	4.61	4.76	.15
Reading DL	31.3	4.79	5.00	.21
Writing ESL	24.8	3.65	4.10	.45
Writing DL	25.9	3.89	4.28	.39
Listening ESL	20.8	4.89	5.07	.18
Listening DL	29.2	5.11	5.21	.10
Speaking ESL	5.7	4.68	4.36	-.32
Speaking DL	9.8	4.78	4.52	-.26

As can be observed in Table 26, the independent variables of pre-test proficiency level, race, gender, and SWD accounted for more variability in academic achievement for students enrolled in a DL program than students enrolled in an ESL program; overall 39% DL and 27.1% ESL, reading 31.3% DL and 28.1% ESL, writing 25.9% DL and 24.8 ESL, listening 29.2% DL and 20.8% ESL, and speaking 9.8% DL and 5.7% ESL. When

comparing mean pre-test proficiency levels, the ESL scores were lower than the DL scores. The same was true for post-test proficiency levels. When comparing mean gain scores students enrolled in the ESL group made larger gains in three out of the five academic areas. In overall academic achievement students in the ESL program had a gain score of .20, on a scale from 1-6, while DL students had a gain score of .17. In writing achievement students in the ESL program had a gain score of .45 while DL students had a gain score of .39. In listening achievement students in the ESL program had a gain score of .18 while DL students had a gain score of.10. Students enrolled in a DL program had a greater gain score only in reading where the score was a .21 while ESL students had a gain score of .15. Both the ESL and DL students had a drop in speaking proficiency, ESL students had a drop of .32 and DL students had a drop of .26.

Summary of the Data

The chapter presented descriptive statistics for the independent and dependent variables, correlation matrices, and stepwise multiple linear regression analyses to analyze the extent that pre-test proficiency level (overall, reading, writing, listening, speaking), gender, race (Black/African American, White), and SWD accounted for academic achievement for students in an ESL program versus a DL program.

The regression analyses showed that for all academic achievement areas SWD was a significant independent variable that accounted for the greatest proportion of the variability and explained 2.6-25.1% of the variability in achievement in ESL programs and 6.6-35.8% of the variability in DL programs. Race (Black/African American) was significant in reading achievement for students enrolled in an ESL program and accounted for 1.7% of the variability. For students enrolled in an ESL instructional

program, the independent variables accounted for 27.1% of the variability in overall achievement, 28.1% of the variability in reading achievement, 24.8% of the variability in writing achievement, 20.8% of the variability in listening achievement, and 5.7% of the variability in speaking achievement. For students enrolled in a DL instructional program, the independent variables accounted for 39% of the variability in overall achievement, 31.3% of the variability in reading achievement, 25.9% of the variability in writing achievement, 29.2% of the variability in listening achievement, and 9.8% of the variability in speaking achievement. Students enrolled in ESL programs had greater gain scores than students enrolled in a DL program in three out of the five academic areas: overall, writing, and listening. DL students had greater gain scores in reading. Both ESL and DL students had a drop in speaking proficiency.

CHAPTER 5

Findings, Discussion, and Implications

This chapter summarizes the findings from the study. The chapter is organized as follows: a summary of the problem, the purpose of the study, the methodology used to conduct the study, a summary of the findings, a discussion of the findings, and implications for practitioners and further research.

Summary of the Problem

ELLs comprise a rapidly growing subgroup within our K-12 student population and their educational needs have become a pressing issue in public schools (Maxwell, 2009). The achievement of ELLs is consistently lower than mainstream native English speakers (Abedi, 2007; Abedi & Dietel, 2004; Fry, 2007). To help this struggling population, the federal government provided funding for instructional programs under Title III. The challenge is to find ways to teach these students a new language as well as academic content. In this era of accountability, ELLs must be able to effectively use English in an academic setting as well as master other academic content to succeed in a public school setting (Coleman & Goldberg, 2010; Francis & Rivera, 2007). There are two major approaches to instructing ELLs: DL and ESL. DL focuses on developing literacy in two languages. ESL focuses on developing literacy only using English. Determining which approach more effectively teaches conversational and academic English as well as content-specific material is a major challenge facing educators of ELLs.

Purpose of the Study

The purpose of the study was to determine the extent pre-test proficiency level, gender, race, and identification as a student with disabilities (SWD) account for the variability in academic achievement for students enrolled in an ESL or a DL instructional program. Analyzing both a traditional English-only ESL program and a DL program addressed one of the main issues cited by researchers—ELL treatments are evaluated in isolation without either a control group or a comparison group (Conger, 2010; Ramirez et al., 1991; Rolstad et al., 2005; Rossell & Baker, 1996; Slavin & Cheung, 2005). The study also employed a pre-test and post-test measure that addressed another criticism in the literature group (Conger, 2010; Ramirez et al., 1991; Rolstad et al., 2005; Rossell & Baker, 1996; Slavin & Cheung, 2005). Many of the studies use only summative data to analyze the effectiveness of an instructional program.

Methodology

The study implemented a quasi-experimental design to determine the effects of pre-test proficiency levels (overall, reading, writing, listening, speaking), gender, race (Black/African American and White), and SWD on the academic achievement of students enrolled in two instructional programs. The DL and ESL instructional programs served as the two independent selection variables while pre-test proficiency levels (overall, reading, writing, listening, speaking), gender, race, and SWD served as independent variables. The dependent variables were the post-test proficiency levels: overall, reading, writing, listening, and speaking.

The treatments used in the study are two popular methods for teaching ELLs that have proven successful in previous studies (Ramirez et al., 1991; Rossell & Baker, 1996).

The ESL program provided for students in this county was well established and used instructional techniques frequently cited (Calderón et al., 2011; Carlo et al., 2005; Collier, 1995; Echeverria, et al., 2006) as effective for ELLs. The DL program in this study applied a 50:50 instructional model for grades K-5 where students spend half the instructional day in English and half in Spanish throughout the length of the program (Christian et al., 2000; Lindholm-Leary, 2004). The DL program in this study was also well designed and followed guiding principles cited by researchers as necessary to develop effective bilingual education programs (Christian et al., 2000; Collier &Thomas, 2001; Lindholm-Leary &Block, 2010).

The sample consisted of fourth- and fifth-grade ELLs enrolled in either ESL or DL programs that were comparable on native language and ethnicity. T-tests demonstrated that the ESL and DL sample mean differences for the independent variables were not statistically significant, and thus the ESL and DL subgroups were comparable in pre-test proficiency level, gender, race, and SWD.

The study used archival data that came from the Assessing Comprehension and Communication in English State to State for English language learners (ACCESS for ELLs). The World-Class Instructional Design and Assessment (WIDA) consortium developed the ACCESS for ELLs by asking their consortia members to review and approve standards that focus on academic English and then taking these agreed upon standards and developing an assessment that measures listening, speaking, reading, and writing (Bauman et al., 2007). The test yields six English language proficiency levels from level one, rating minimal proficiency, to level six, rating maximum proficiency, in five categories: overall proficiency, reading proficiency, writing proficiency, listening

proficiency, and speaking proficiency. The ACCESS was chosen because it was a test designed for ELLs, so it should more appropriately assess student achievement than a large scale state mandated test designed for mainstream native English speaking students.

Descriptive statistics including minimum, maximum, mean, standard deviation, and variance were provided for all the independent and dependent variables. Stepwise multiple linear regression analyses were utilized to determine to what extent pre-test proficiency level, gender, race, or SWD accounted for variability in overall academic achievement, reading academic achievement, writing academic achievement, listening academic achievement, and speaking academic achievement of ELLs enrolled in an ESL or DL instructional program.

Summary of the Findings

This section summarizes the descriptive statistics and the study findings by research question.

Descriptive statistics were presented by the total sample and by instructional program. The minimum and maximum pre-test and post-test proficiency levels range from 1.0 to 6.0. The total sample mean pre-test proficiency levels ranged from 3.71-4.96 while the total sample post-test means ranged from 4.15-5.11 (Table 3). The mean for pre-test overall proficiency level was 4.40 while the post-test overall proficiency level was 4.59, a mean gain of .19 (Table 3). For reading the pre-test proficiency level mean was 4.67 while the post-test proficiency level mean was 4.83, a mean gain of .16 (Table 3). For writing the pre-test proficiency level mean was 3.73 while the post-test proficiency level mean was 4.15, a mean gain of .42 (Table 3) which was the largest gain for any of the proficiency levels studied. For listening the pre-test proficiency level mean

was 4.96 while the post-test proficiency level mean was 5.11, a mean gain of .15 (Table 3). For speaking the pre-test proficiency level mean was 4.71 while the post-test proficiency level mean was 4.40, a mean drop of .31 (Table 3). This was the only academic area studied where there was a drop in proficiency. When reviewing the means by instructional program, the ESL sample means were lower than the total sample means in all of the achievement categories. The pre-test means for ESL were lower than the total sample means by a range of 03-.08 (Table 4). The post-test means for ESL were lower than the total sample means by a range of .04-.07 (Table 4). The DL means were higher than the total sample means in all achievement categories. The pre-test means for DL were higher than the total sample means by a range of .07-.20 (Table 5). The post-test means were higher than the total sample means by a range of .10-.18 (Table 5).

Stepwise multiple regressions were used to answer the research questions. *Question 1: To what extent does pre-test proficiency level, gender, race, or SWD explain variability of overall academic achievement for students enrolled in an ESL program versus students enrolled in a DL program?* To analyze the relationships among the independent variables the researcher produced a correlation matrix using the Pearson correlation and conducted a stepwise multiple regression to answer the research question. SWD was a significant variable for students enrolled in both the ESL and DL programs. SWD accounted for 25.1% of variability in overall academic achievement for ESL students and 35.8% of variability in overall academic achievement for DL students. All the independent variables accounted for 27.1% of variability in overall achievement for ESL students and 39% of the variability in overall achievement for DL students. For students enrolled in an ESL program the independent variables of pre-test proficiency

level, gender, race, and SWD had less impact on overall achievement than they did for students enrolled in a DL program.

Question 2: To what extent does pre-test proficiency level, gender, race, or SWD explain variability of overall reading achievement for students enrolled in an ESL program versus students enrolled in a DL program? To analyze the relationships among the independent variables the researcher produced a correlation matrix using the Pearson correlation and conducted a stepwise multiple regression to answer the question. SWD was a significant variable for students enrolled in both the ESL and DL programs. SWD accounted for 25% of the variability in reading achievement for ESL students and 27.3% of the reading achievement for DL students. Race (Black/African American) was also a significant variable for students enrolled in an ESL program accounting for 1.7% of the variability in reading achievement. All the independent variables accounted for 28.1% of the variability in reading achievement for ESL students and 31.3% of the variability in reading achievement for DL students. For students enrolled in an ESL program the independent variables of pre-test proficiency level, gender, race, and SWD had less impact on reading achievement than they did for students enrolled in a DL program.

Question 3: To what extent does pre-test proficiency level, gender, race, or SWD explain variability of overall writing achievement for students enrolled in an ESL program versus students enrolled in a DL program? To analyze the relationships among the independent variables the researcher produced a correlation matrix using the Pearson correlation and conducted a stepwise multiple regression to answer the research question. SWD was a significant variable for students enrolled in both the ESL and DL instructional programs. SWD accounted for 23.3% of the variability in writing achievement for ESL students and

23.6% of the writing achievement for DL students. All of the independent variables accounted for 24.8% of the variability in writing achievement for ESL students and 25.9% of the variability for DL students. For students enrolled in an ESL program the independent variables of pre-test proficiency level, gender, race, and SWD had less impact on writing achievement than they did for students enrolled in a DL program.

Question 4: To what extent does pre-test proficiency level, gender, race, or SWD explain variability of overall listening achievement for students enrolled in an ESL program versus students enrolled in a DL program? To analyze the relationships among the independent variables the researcher produced a correlation matrix using the Pearson correlation and conducted a stepwise multiple regression to answer the research question. SWD was a significant variable for students enrolled in both the ESL and DL programs. SWD accounted for 18.7% of the variability in listening achievement for ESL students and 25.5% of the variability in reading achievement for the DL students. All the independent variables accounted for 20.8% of the variability in listening achievement for ESL students and 29.2% for DL students. For students enrolled in an ESL program the independent variables of pre-test proficiency level, gender, race, and SWD had less impact on listening achievement than they did for students enrolled in a DL program.

Question 5: To what extent does pre-test proficiency level, gender, race, or SWD explain variability of overall speaking achievement for students enrolled in an ESL program versus students enrolled in a DL program? To analyze the relationships among the independent variables the researcher produced a correlation matrix using the Pearson correlation and conducted a stepwise multiple regression to answer the research question. SWD was a significant variable for students enrolled in the ESL and DL programs. SWD

accounted for 2.6% of the variability in speaking achievement for ESL students and 6.6% of the variability in listening achievement for DL students. All the independent variables accounted for 5.7% of the variability in speaking achievement for ESL students and 9.8% of the variability in speaking achievement for DL students. For students enrolled in an ESL program the independent variables of pre-test proficiency level, gender, race, and SWD had less impact on listening achievement than they did for students enrolled in a DL program.

Discussion

Stepwise multiple regression analyses were conducted to determine to what extent pre-test proficiency level, gender, race (White, Black/African American), and SWD accounted for the variability in academic achievement for students enrolled in and ESL or DL instructional program. The ESL and DL programs were the independent selection variables used to compare the variability in achievement for students enrolled in either program. In all five achievement areas studied (overall, reading, writing, listening, and speaking), SWD was a significant variable accounting for the majority of variability in achievement. Race (Black/African American) was the only other significant independent variable accounting for 1.7% of the variability in reading achievement for students enrolled in an ESL program. SWD accounted for variability of 2.6-35.8% (Tables 8, 9, 12, 13, 20, 21, 24, 25) and had a negative correlation to post-test proficiency level in all achievement areas. Therefore, SWD accounted for a significant variability in underachievement of ELLs. This finding is in line with other researchers who cite the complexity in properly identifying learning disabilities in ELLs as a cause for the

underachievement of ELLs in school (Abedi, 2009; Anderson et al., 2005; Artiles et al., 2005).

Of the total sample about 33% of students were identified as SWD (Table 3). In the total county fourth- and fifth-grade population, about 15% of the students were identified as SWD; so the ELL SWD percentage is about twice the average of the total population (Virginia Department of Education, 2010). Artiles et al. found overrepresentation of ELLs in special education programs in their 2005 study and indicated that it may be due to the lack of English proficiency causing students to perform poorly on measures to identify learning disabilities. While this study cannot determine whether the overrepresentation of students with disabilities was due to their lack of English proficiency, it is important to note the percentage of students with a disability is twice the number identified in the total student population.

The percentage of variability accounted for by SWD ranged from 2.6-35.8% (Tables Tables 8, 9, 12, 13, 20, 21, 24, 25) and in all cases was the significant variable accounting for achievement. This amount of variability is important since researchers cite the lack of effective ways to teach ELLs with learning disabilities as an important factor contributing to the achievement gap (Sáenz et al., 2005). The findings support the need to find effective ways to help ELLs improve their achievement, especially if they are also identified as students with a disability.

Another independent variable expected to account for a significant percentage of the variability was the pre-test proficiency level. However, the pre-test level accounted for very small percentages in variability ranging from .1-2% and was not a significant variable in any of the academic areas. Pre-tests are used to strengthen a study design

since they can provide a baseline and are usually the variable most highly correlated to the post-test score (Trochim, 2005). The findings of the study suggest that the pre-test/post-test correlation may have been attenuated due to the homogeneity of the sample (Cole et al., 2011). The descriptive statistics in Tables 3-5 showed that small gains were made by the sample and small gains combined with a homogenous sample can cause relatively large changes in relative position within the sample which weakened the correlation between the pre-test and post-test proficiency levels (Cole et al., 2011).

The two instructional programs were used as the independent selection variables to determine differences in variability for students enrolled in either ESL or DL programs. For students enrolled in the DL program the independent variables accounted for a larger percentage of variability than for students enrolled in the ESL program. The findings suggest that the ESL instructional program had more of an impact on academic achievement than the DL program. The mean gain scores in Table 26 also suggest that the ESL program displayed greater results than the DL program. The ESL program outscored the DL program in overall achievement by .03, in writing by .06, and in listening by .08 (Table 26). The DL program outscored the ESL program only in reading by .06 (Table 26). Both the ESL and DL programs had drops in speaking proficiency. The study findings suggest that less variability and greater gains were made by students enrolled in the ESL program when compared to students enrolled in the DL program and support research on the effectiveness of structured English-only instruction that integrates language and content skills as a way to help students be successful in mainstream classes (Conger, 2010; Rossell, 2003; Rossell 2005; Short et al. 2011). Intense practice in English may have contributed to the greater gain ESL students demonstrated in the study.

Since the higher WIDA levels address the finer skills of English language proficiency, increased time in English would allow students to practice and refine the sophisticated skills of complex language output as described in Figure 2.

However, both groups had small gains in proficiency. In terms of gain scores, the descriptive statistics (Table 3) showed that all students, regardless of program, had small gains in four of the academic areas: overall proficiency level rose by .19, reading proficiency level rose by .16, listening proficiency level rose by .15, and writing proficiency level rose by .42, the biggest gain in achievement. Speaking proficiency level actually decreased by .31. The gains were small but that may be attributed to how high the proficiency levels were already. The mean pre-test proficiency levels ranged from 3.73-4.96 (Table 3). The county policy for the study sample reclassifies students as English proficient when they reach WIDA level 5. Many of these students were in WIDA level 4 at the beginning of the year. This meant that they had some academic language in other content areas, could produce sentences with varying linguistic complexity, and had minimal phonological, syntactic, or semantic errors in their oral or written language (Figure 2). The WIDA level 5 characteristics differ slightly to the level 4 criteria and focus on a more consistent demonstration of the academic language, variety of sentence lengths and complexities in longer writing pieces, and oral and written language approaching comparability to a native English speaker at the same grade level (Figure 2). The differences between WIDA levels 4-5 are more difficult to attain since they address the final skills needed to achieve English proficiency, so the gains will not be as large as they are when moving up the lowest levels of proficiency. The study data support that

there were small gains made by students whether in the ESL or DL program, but students in the ESL program made greater gains than students in the DL program.

Implications for Practitioners

Several of the study findings provide areas for practitioners to focus on as they strive to develop educational programs to meet the needs of ELLs. The regression analyses consistently showed SWD as a significant variable accounting for 2.6-35.8% of the variability in achievement (Tables 8, 9, 12, 13, 20, 21, 24, 25). The descriptive statistics showed that about 31% of the sample was made up of students with disabilities while 15% of the whole population was made up of students with disabilities. The large percentage of variability accounted for by SWD along with the fact that almost a third of the sample was identified as having a learning disability, suggests that schools need to focus on this population and how to best teach them English.

Schools need to implement effective ways to assess whether ELL underachievement is due to lack of English proficiency or a learning disability (Abedi, 2009; Anderson et al, 2005; Artiles et al, 2005; Minnema et al, 2005). Once students are identified, schools need to determine the best program to address the needs of ELLs with learning disabilities. Sáenz et al. (2005) advocate the use of peer tutors and cooperative learning to help ELLs with disabilities succeed in the classroom. The DL program in the study employed peer tutoring as a main part of their program objectives and also strived for a diverse student body including students with special needs (Forbes-Ullrich & Perdomo, 2005). However, the data show that the ESL program demonstrated greater gains than the DL program in all but one academic area, reading. The study findings suggest that schools which have substantial ELL populations, especially with learning

disabilities, should explore the implementation of an ESL program to serve that population since it is shown to account for greater gains in overall, writing, and listening achievement.

The study findings revealed that speaking proficiency dropped for all students in the sample by .31 (Table 3). For students enrolled in an ESL program proficiency dropped by .32 and for students enrolled in a DL program it dropped by .26. The findings suggest that schools with instructional programs for ELLs need to review the curriculum and instruction in the area of speaking. In this area of accountability, schools are focused on reading, writing, and math and that may have contributed to the fact that the instructional program accounted for no variability in listening and speaking achievement. With the move to new national standards as part of the reauthorization of the Elementary and Secondary Education Act, schools will need to demonstrate students are proficient in oral (listening and speaking) communication. The findings from the study may provide some justification for reviewing the curriculum and implementing some instructional activities that more effectively teach listening and speaking skills.

The study findings indicated that students enrolled in an ESL program made greater gains than students in a DL program for three out of the five academic areas studied. The proficiency levels for students enrolled in the ESL program increased from .18-.45 while the proficiency levels for students enrolled in the DL program increased .10-.39. Schools with limited resources that may only be able to support one English language program may consider the findings of the study when selecting their ELL instructional program. Schools with more ample resources may look at the study findings by academic area. Students in the ESL program made greater gains in overall proficiency,

writing proficiency, and listening proficiency. Students in the DL program made greater gains in reading proficiency. Schools with more resources may consider the study findings and implement mixed models based on the gain scores. Another finding to note is that the students in the study were also at the highest levels of proficiency before being reclassified, so for that type of population the findings indicated that the ESL program provided greater gains in proficiency for a majority of the academic areas studied.

Implications for Further Study

Several of the study findings provide areas for further study. The data show that students enrolled in an ESL program made greater gains in a school year than did students in the DL program. Students enrolled in the ESL program showed greater gains in overall, writing, and listening proficiency. The study sample was comprised of students who were at the highest levels of proficiency before being reclassified as proficient. Further studies should be conducted to explore if students at the lower levels of proficiency benefit from ESL instruction at the same rate as students at the highest levels. The data also showed that DL students had a higher pre-test proficiency level and post-test proficiency level even though these students did not make as great of a gain throughout the year. Further studies of DL students at lower proficiency levels should be conducted to explore if the DL program is more effective for students at these lower proficiency levels. The findings from the study support the effectiveness of English-only ELL instruction, and should be replicated to further demonstrate if the ESL model implemented in this county is more effective than the DL instructional model.

The findings suggest little correlation between the pre-test and post-test measures which is unusual. Pre-tests are used because of their high correlation to the post-test

(Trochim, 2005). Correlation matrices showed weak correlations to the post-test proficiency level ranging from -.016-.168 (Tables 6, 7, 10, 11, 14, 15, 18, 19, 22, 23). The findings of the study suggest that the pre-test/post-test correlation may have been attenuated due to the homogeneity of the sample (Cole et al., 2011). Further study regarding the attenuation of the pre-test/post-test correlation may provide reasons for the weakened correlations that could be accounted for in future study designs. There could also be further studies to determine other baseline measures that could be used with a homogenous sample that provide a high correlation to the post-test measure as well as a way to baseline a sample.

The data suggest that the selection of a standardized assessment rather a measure especially designed for the study as a pre-test/post-test measure provided may have also impacted the pre-test/post-test correlation. The regression analyses found the pre-test accounted for very small percentages in academic achievement no matter what the instructional program and never accounting for significant variability. For overall achievement, the pre-test accounted for .6% of the variability for ESL students and .1% for DL students; reading achievement .4% for ESL students and .3% for DL students; writing achievement .1% for ESL students and .8% for DL students; listening 1.2% for ESL students and .2% for DL students; and for speaking 2.0% for ESL students and .1 for DL students. The findings support research stating that using a standardized assessment versus an assessment specifically designed to measure the intervention may underestimate the effect of the treatment (Olsen et al., 2011). There could be further studies that use assessment measures specifically designed to measure treatments that

could better determine the effects the instructional programs have on academic achievement.

Descriptive statistics for this population showed that students were already at the upper levels of the proficiency scale with pre-test means ranging from 3.73-4.96 (Table 3). Many of the students had a pre-test proficiency level of 4, one level below reclassification as English proficient according to the policy of the county from which the sample was acquired. According to the WIDA levels, the difference between level 4 and level 5 is the focus on a more consistent demonstration of the academic language, variety of sentence lengths and complexities in longer writing pieces, and oral and written language approaching comparability to a native English speaker at the same grade level (Figure 2). Even though students showed gains in four out of the five academic areas, the gains were small ranging from .15-.19 and there was a .31 drop in the area of speaking. The small gains reported in the findings suggest that exploring ways to measure discreet skills could more effectively assess the impact of the program on academic achievement.

A significant variable accounting for the majority of academic achievement in all areas was SWD; the variable accounted for 2.6-35.8% of the variability in achievement which was a majority of the achievement accounted for by all the variables included in the regression analyses. The variability accounted for by the other independent variables was minimal compared to the variability accounted for by SWD. Further studies on how to better address the needs of ELLs with disabilities could help change the methods used in instructional programs so they can effectively teach ELLs with and without learning disabilities.

The variables in the regression analyses were able to account for less than 40% of the variability in ELL achievement. In overall achievement, the variables accounted for 29.6% of the variability in achievement for ESL students and 39% of the variability in achievement for DL students (Table 26). In reading achievement, the variables accounted for 28.1% of the variability in achievement for ESL students and 31.3% of the variability in DL students (Table 26). For writing achievement, the variables accounted for 24.8% of the variability in achievement for ESL students and 25.9% of the variability for DL students (Table 26). In listening achievement, the variables accounted for 20.8% of the variability in achievement for ESL students and 29.2% of the variability for DL students (Table 26). In speaking achievement, the variables accounted for 5.7% of the variability in achievement for ESL students and 9.8% of the variability in DL students (Table 26). The variables used were pre-test proficiency level, gender, race (White, Black/African American), and SWD. The treatment variables used as independent selection variables were the DL and ESL instructional programs. Further studies that select different variables for achievement would help determine what other factors account for ELL achievement. Some of the variables that could be studied could be length of time in the program since researchers have found the longer students are in an English instructional program the less able they are to catch up to their native English language speaking peers (Kim & Herman, 2009; Rossell, 2005).

Conclusion

This study sought to add to the body of research on effective programs to instruct ELLs in English as well as other academic areas. ELLs are a significant part of the school aged population comprising about 10% of the total population (NCES, 2009) and the

population has increased at a rate of 57 % versus 4% for all students since 1995 (NCELA, 2008). This population continues to struggle when compared to their native English speaking peers as shown by a consistent achievement gap in math and reading (Abedi & Dietel, 2004; Fry, 2007). The purpose of the study was to determine the effects of pre-test proficiency level, race, gender, and SWD on academic achievement in five areas (overall, reading, writing, listening, and speaking) for students enrolled in an ESL or a DL instructional program. SWD was a significant variable in all academic areas accounting for the greatest percentage of variability in student achievement. The data showed that students enrolled in an ESL program made greater gains than their DL peers in overall proficiency, writing proficiency, and listening proficiency. DL students made greater gains in reading proficiency. The study findings highlight the need to continue exploring how to effectively educate ELLs, especially those with learning disabilities, to acquire the skill set to succeed in our public schools.

References

Abedi, J. (2007). *English language proficiency assessment in the nation: Current status and future practice.* University of California, Davis: The Regents of the University of California.

Abedi, J. (2009). English language learners with disabilities: Classification, assessment, and accommodation issues. *Journal of Applied Testing Technology*, *10*(2), 1-30.

Abedi, J., Dietel, J. N., & California Univ., L. N. (2004). *CRESST Policy Brief 7. Winter 2004.* Los Angeles, CA: National Center for Research on Evaluation Standards and Student Testing.

Alanís, I., & Rodriguez, M. (2008). Sustaining a dual language immersion program: Features of success. *Journal of Latinos & Education, 7*, 305-319.

Anderson, M., Minnema, J., Thurlow, M., & Hall-Lande, J. (2005). *Confronting the unique challenges of including English language learners with disabilities in statewide assessments* (ELLs with Disabilities Report 9). Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes.

Artiles, A. J., Rueda, R., Salazar, J., & Higareda, I. (2005). Within-group diversity in minority disproportionate representation: English language learners in urban school districts. *Exceptional Children*, *71*(3), 283-300.

Aud, S., Fox, M., & KewalRamani, A. (2010). *Status and trends in the education of racial and ethnic groups* (NCES 2010-015). Washington, DC: U.S. Government Printing Office.

August, D., & Shanahan, T. (2006). *Developing literacy in second-language learners: Report of the national literacy panel on language minority children and youth.* Mahwah, NJ: Lawrence Erlbaum Associates.

Baker, C. (2001). An introduction to bilingual education. In C. Baker (Ed.), *Foundations of Bilingual Education and Bilingualism* (pp. 182-202). Clevedon, UK: Multilingual Matters Ltd.

Baker, K. (1998). Structured English immersion: Breakthrough in teaching limited-English-proficient students. *Phi Delta Kappan, 80,* 199-204.

Bauman, J., Boals, T., Cranley, E., Gottleib, M., & Kenyon, D. (2007). Assessing comprehension and communication in English state to state for English language learners (ACCESS for ELLs). In Abedi, J. (Ed.*), English Language Proficiency Assessment in the Nation: Current Status and Future Practice* (pp.81-91). University of California, Davis: The Regents of the University of California.

Boyle, A., Taylor, J., Hurlburt, S., & Songa K. (2010) *Title III accountability: Behind the numbers. ESEA evaluation brief: The English language acquisition, language enhancement, and academic achievement act.* Washington, D.C.: US Dept of Ed publication.

Brisk, M. E., & Jeffries, J. (2008). Academic English. In González, J.M. (Ed.), *Encyclopedia of Bilingual Education* (pp.1-4). Los Angeles, CA: Sage Publications.

Calderón, M. E., Hertz-Lazarowitz, R., & Slavin, R. E.(1998). Effects of bilingual cooperative integrated reading and composition on students making the transition from Spanish to English reading. *Elementary School Journal, 99,*153-165.

Calderón, M., Slavin, R., & Sánchez, M. (2011). Effective instruction for English learners. *Future of Children, 21*(1), 103-127.

Callahan, R., Wilkinson, L., & Muller, C. (2010). Academic achievement and course taking among language minority youth in U.S. schools: Effects of ESL placement. *Educational Evaluation and Policy Analysis, 32*(1), 84-117.

Carlo, M. S., August, D., Mclaughlin, B., Snow, C., Dressler, C., Lippman, D., & White, C. E. (2008). Closing the gap: Addressing the vocabulary needs of English-language learners in bilingual and mainstream classrooms. *Journal of Education, 189*(1/2), 57-76.

Chen, C. S. (2011). *Public elementary and secondary school student enrollment and staff counts from the Common Core of Data: School year 2009–10* (Report No. 2011-347). Washington, D.C.: U.S. Department of Education.

Christian, D., Howard, E., & Loeb, M. (2000). Bilingualism for all: Two-way immersion education in the United States. *Theory into Practice, 39,* 258-266.

Cobb, B., Vega, D., & Kronauge, C. (2006). Effects of an elementary dual language immersion school program on junior high school achievement. *Middle Grades Research Journal*, *1*(1), 27-47.

Cole, R., Haimson, J., Perez-Johnson, I., & May, H. (2011). *Variability in pretest-posttest correlation coefficients by student achievement level*. NCEE Reference Report 2011-4033. Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, 2011.

Coleman, R., & Goldenberg, C. (2010). What does research say about effective practices for English learners? Part II: Academic language proficiency. *Kappa Delta Pi Record*, *46*(2), 60-65.

Collier, V. P. (1995). *Promoting Academic Success for ESL Students: Understanding Second Language Acquisition.* Woodside, New York: Bastos Educational Publications.

Conger, D. (2010). Does bilingual education interfere with English-language acquisition?. *Social Science Quarterly (Blackwell Publishing Limited)*, *91*, 1103-1122. doi:10.1111/j.1540-6237.2010.00751.x

Cummins, J. (1984). Wanted: A theoretical framework for relating language proficiency to academic achievement among bilingual students. In C. Rivera (Ed.), *Language proficiency and academic achievement*,(pp. 2-19). Clevedon, Avon, England: Multilingual Matters

Cummings, J. (1992). Bilingual education and English immersion: The Ramírez report in theoretical perspective. *Bilingual Research Journal, 16*, 91-104.

De Jesús, S. C. (2008). An astounding treasure: Dual language education in a public school setting. *Centro Journal, 20*, 192-217.

Echevarria, J., Short, D., & Powers, K. (2006). School reform and standards-based education: A model for English-language learners. *Journal of Educational Research*, *99*, 195-210.

Elementary & Secondary Education Act of 2001, Title III § 3102. Retrieved from http://www2.ed.gov/policy/elsec/leg/esea02/pg40.html.

Edvantia (2007). Research review: What research says about preparing English language learners for academic success. Retrieved from http://www.centerforpubliceducation.org/site/c.kjJXJ5MPIwE/b.3531983/k.A79C/Preparing_English_language_learners_for_academic_success.htm

Forbes-Ullrich, M. A., & Perdomo, M. (2005). Two-way immersion program evaluation report. Arlington, VA: Arlington Public Schools.

Francis, D. F, & Rivera, M. O. (2007). Principles underlying English language proficiency tests and academic accountability of ELLs. In Abedi, J. (Ed.*), English Language Proficiency Assessment in the Nation: Current Status and Future Practice* (pp.13-32). University of California, Davis: The Regents of the University of California.

Fry, R., & Pew Hispanic, C. (2007). *How Far behind in Math and Reading Are English Language Learners?* Report. Washington, D.C.: Pew Hispanic Center.

Genesee, F., Lindholm-Leary, K., Saunders, W., & Christian, D. (2005). English language learners in U.S. schools: An overview of research findings. *Journal of Education for Students Placed at Risk, 10*, 363-386.

Hansen-Thomas, H. (2008). Sheltered instruction: Best practices for ELLs in the mainstream. *Kappa Delta Pi Record, Summer 2008,*165-169.

Hemphill, F., Vanneman, A., & National Center for Education Statistics (2011). *Achievement gaps: How Hispanic and White students in public schools perform in mathematics and reading on the National Assessment of Educational Progress.* Statistical Analysis Report. NCES 2011-459. Washington, D.C.: National Center For Education Statistics.

Howard, E., Christian, D., & Center for Research on Education, D. (2002). *Two-way immersion 101: Designing and implementing a two-way immersion education program at the elementary level.* Educational Practice Report. Retrieved from ERIC database. ED 473 082.

Kim, J. (2011). *Relationships among and between ELL status, demographic characteristics, enrollment history, and school persistence* (CRESST Report 810). Los Angeles, CA: University of California, National Center for Research on Evaluation, Standards, and Student Testing (CRESST).

Kim, J., & Herman, J. L. (2009). A three-state study of English learner progress. *Educational Assessment, 14,* 212-231. doi:10.1080/10627190903422831.

Krashen, S. (2008). Language education: Past, present and future. *RELC Journal, 39,* 178-187.

Lindholm-Leary, K., & Block, N. (2010). Achievement in predominantly low SES/Hispanic dual language schools. *International Journal of Bilingual Education and Bilingualism, 13*(1), 43-60.

Lindholm-Leary, K. J., & Genesee, F. (2010). Alternative educational programs for English language learners. In California Department of Education (Ed.), *Research on English language learners,* (pp. 323-382). Sacramento: CDE Press.

Lindholm-Leary, K., & Hernandez, A. (2011). Achievement and language proficiency of Latino students in dual language programmes: Native English speakers, fluent English/previous ELLs, and current ELLs. *Journal of Multilingual and Multicultural Development, 32,* 531-545.

Lindholm-Leary, K. J., & Howard, E. (2008). Language and academic achievement in two-way immersion programs. In T. Fortune and D. Tcdick (Eds.), *Pathways 10 bilingualism: Evolving perspectives on immersion education,* (pp.177-200). Clevedon. U.K.: Multilingual Matters.

Lindholm-Leary, K.J. 2001. *Dual language education.* Clevedon, U.K.: Multilingual Matters.

Office of English Language Acquisition, Language Enhancement, and Academic Achievement for Limited English Proficient Students (2006). *Biennial Report to Congress on the Implementation of the Title III State Formula Grant Program, School Years 2004–06.* Washington, D.C. Retrieved from: http://www.ncela.gwu.edu/files/uploads/3/Biennial_Report_0406.pdf.

Olsen, R. B., Unlu, F., Jaciw, A. P., & Price, C. (2011). *Estimating the Impacts of Educational Interventions Using States Tests or Study-Administered Tests.* (NCEE 2012-4016). Washington, D.C.: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

Maxwell, L. A. (2009, January 8). Shifting landscape: Immigration transforms communities. *Education Week,* 28(17), 1-2.

Menken, K. (2010). NCLB and English language learners: Challenges and consequences. *Theory Into Practice*, *49*(2), 121-128. doi:10.1080/00405841003626619.

Minnema, J., Thurlow, M., Anderson, M., & Stone, K. (2005). *English language learners with disabilities and large-scale assessments: What the literature can tell us* (ELLs with Disabilities Report 6). Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes.

Mitchell, R. & Myles, F. (2004). *Second language learning theories*. New York: Oxford University Press.

National Center for Education Statistics. (2009). Figure 6-2: Percentage of children who spoke a language other than English at home and spoke English with difficult, by state and jurisdiction: 2009. Retrieved from http://nces.ed.gov/programs/coe/figures/figure-lsm-2.asp.

National Center for Education Statistics, (2011*). Achievement gaps: How Hispanic and White students in public schools perform in mathematics and reading on the National Assessment of Educational Progress. Highlights.* NCES 2011-485. Washington, D.C.: National Center For Education Statistics.

National Clearinghouse for English Language Acquisition. How many school-aged limited English proficient (LEP) students are there in the U.S.? Retrieved from http://www.ncela.gwu.edu/faqs/view/4.

Norusis, M. J. (2005). SPSS 14.0 Guide to Data Analysis. Upper Saddle River, NJ: Prentice Hall.

Ramirez, J. D. (1992). Executive summary of the longitudinal study of immersion strategy, early-exit and late-exit transitional bilingual education programs for language-minority children. *Bilingual Research Journal, 16*, 2-62.

Ramirez, J. D., Yuen, S. D., Ramey, D. R., Pasta, D. J., & Billings, D.(1991). *Final report: Longitudinal study of immersion strategy, early-exit and late-exit transitional bilingual education programs for language-minority children.* San Mateo, CA: Aguirre International.

Rolstad, K., Mahoney, K., & Glass, G. V. (2005). The big picture: A meta-analysis of program effectiveness research on English language learners. *Educational Policy, 19*, 572-594.

Rolstad, K., Mahoney, K., & Glass, G. V. (2008). The big picture in bilingual education: A meta-analysis corrected for Gersten's coding error. *Journal of Educational Research & Policy Studies, 8*(2), 1-15.

Rossell, C. H., & Baker, K. (1996). The educational effectiveness of bilingual education. *Research in The Teaching of English, 30*, 307-374.

Rossell, C. H., & ERIC Clearinghouse on Urban Education, N. Y. (2003). *Policy matters in teaching English language learners: New York and California.* Urban Diversity Series.

Rossell, C. (2005). The flawed requirements for limited English proficient children of the No Child Left Behind Act. *Journal of Education, 186*, 29-40.

Sáenz, L. M., Fuchs, L. S, Fuchs, D. (2005) Peer-assisted learning strategies for English language learners with learning disabilities. *Exceptional Children, 71*, 231-247.

Saunders, B., & Goldenberg, C. (1996). Four primary teachers work to define constructivism and teacher-directed learning: Implications for teacher assessment. *Elementary School Journal, 97*, 139–61.

Slavin, R. E., & Cheung, A. (2005). A synthesis of research on language of reading instruction for English language learners. *Review of Educational Research, 75,* 247-284.

Short, D. J., Echevarria, J., & Richards-Tutor, C. (2011). Research on academic literacy development in sheltered instruction classrooms. *Language Teaching Research, 15,* 363-380.

Smith, J., Coggins, C., & Cardoso, J. (2008). Best practices for English language learners in Massachusetts: Five years after the Question 2 mandate. *Equity & Excellence in Education, 41,* 293-310.

Snow, C. E, & Biancarosa, G. (2003). Adolescent literacy and the achievement gap: What do we know and where do we go from here?. Retrieved from http://www.cgcs.org/meetings/snow.pdf

Tong, F., Irby, B. J., Lara-Alecio, R., & Mathes, P. G. (2008). English and Spanish acquisition by Hispanic second graders in developmental bilingual programs: A 3-year longitudinal randomized study. *Hispanic Journal of Behavioral Sciences, 30,* 500-529.

Thomas, W. P. & Collier,V. (1997). *School effectiveness for language minority students*. Washington D.C.: National Clearinghousefor Bilingual Education.

Thomas, W. P. & Collier,V. (2001). *A national study of school effectiveness for language minority students' long-term academic achievement, final report, project 1.1.* Retrieved from http://www.crede.ucsc.edu/research/llaa/1.1_final.html

Trochim, W. M. K. (2005). Research methods: The concise knowledge base. Mason, OH: Thomson.

U.S. Department of Education (2003). *Report Cards Title I, Part A: Non-Regulatory Guidance.* Washington, DC: U.S. Department of Education.

Valverde, L. A., & Armendáriz, G. J. (1999). Important administrative tasks resulting from understanding bilingual program designs. *Bilingual Research Journal, 23*(1), 1-10.

Viadero, D. (2009, January 8). Delving deep: Research hones focus on ELLs. *Education Week*, 28(17), 3-6.

Violand-Sanchez, E., Tabatabai, F., Van Horne, R., & Varela, E. (2007). Evaluation report: Elementary English for speakers of other languages/high intensity language training (ESOL/HILT). Arlington, VA: Arlington Public Schools.

Virginia Department of Education. (2010). *Virginia State Report Card*. Retrieved from https://p1pe.doe.virginia.gov/reportcard/report.do?division=All&schoolName=All

Virginia Department of Education (2010). *Fall Membership.* Retrieved from http://bi.vita.virginia.gov/doe_bi/rdPage.aspx?rdReport=Main&subRptName=Fallmembership