A COMPARATIVE STUDY OF STUDENT RESULTS OF TWO ENGLISH AS A SECOND LANGUAGE PROGRAMS

BY

YVETTE LOZANSKI

Dissertation Committee

Dr. Edward Izbicki Sr., Mentor
Dr. Maureen Blue, Committee Member
Dr. David Turi, Committee Member

Submitted in Partial Fulfillment
Of the Requirements for the Ed.D. Degree
Doctor of Education in Educational Leadership
Saint Peter’s University

May, 2016
Abstract

Educators seek the best methods to teach their students on a daily basis. ELL learners not only have the hard task of acculturation, they are also expected to reach the same level of proficiency as their native English-speaking peers. Researchers have made efforts to seek out the best type of instruction for these ELL students. Accordingly, the debate continues about what type of instruction—inclusion or exclusion—best benefits these types of learners. This study was designed to examine the effect the types of instruction has on the ELL learner. The WIDA ACCESS test was administered to students at the end of the school year after receiving either inclusion or exclusion instruction to grades kindergarten through three in two districts in New Jersey. The outcomes of the scores were then analyzed using the independent samples t-test in a quantitative study design. Findings showed that in the four domains of listening, speaking, reading, and writing, only certain areas showed to be significant that the type of instruction has an impact on the ELL learner. This study confirmed that the delivery of instruction could play a significant role in the education of ELLs.
ACKNOWLEDGEMENTS

I would not have been able to finish my dissertation without my committee members. My deepest gratitude is to my mentor, Dr. Edward Izbicki. I have been fortunate to have a mentor who knew when to push me and knew when to give me freedom to work on my own. His expertise, genius, guidance, support, and patience have brought me to this point in my life and, for that, I will be forever grateful.

Dr. Blue, if it was not for you, I would not be here at all. I would like to express my gratitude to you. You allowed me into your program and into your life. I appreciate everything you have done and the amount of patience you have had receiving the obscene amount of emails I have tortured you with.

I would also like to thank Dr. Turi. Your knowledge has taught me the skills I needed in order to complete this dissertation. I appreciate the time and effort you put into helping me, including putting up with the extreme number of office visits.

I would also like to thank all of those people who have helped me directly or indirectly, and who have lent a helping hand, including Ms. Bonnie Luberto and Ms. Geri Ledford. Mom, for giving us everything when we should have had nothing. Finally, to Tommy who would not let me quit.
# Table of Contents

Abstract.................................................................................................................................................... iii
Acknowledgements........................................................................................................................................ iv
Chapter 1 INTRODUCTION......................................................................................................................... 1
Definition of the Terms.................................................................................................................................. 3
Chapter 2 LITERATURE REVIEW ............................................................................................................... 10
Chapter 3 METHODOLOGY.......................................................................................................................... 20
Setting and Sample ...................................................................................................................................... 22
Instrumentation and Resources .................................................................................................................... 23
Data Collection ............................................................................................................................................. 25
Null Hypothesis Alternative Hypothesis ....................................................................................................... 26
Assumptions, Limitations, Delimitations ....................................................................................................... 28
Chapter 4 Results ......................................................................................................................................... 29
Demographics ............................................................................................................................................... 29
Kindergarten .............................................................................................................................................. 30
Grade 1 ....................................................................................................................................................... 37
Grade 2 ....................................................................................................................................................... 44
Grade 3 ....................................................................................................................................................... 50
Summary ..................................................................................................................................................... 57
Chapter 5 INTERPRETATION AND RECOMMENDATIONS ..................................................................... 58
Overview of the Problem ............................................................................................................................... 58
Purpose of the Study/Research Questions ................................................................................................. 58
List of Tables

Table 1. Independent Sample T-Test Between Students Receiving Pull-out Resources and Push-in Resources ...............................................................31
Table 2. Independent Sample T-Test between students receiving Pull-out resources and Push-In resources ................................................................................................................32
Table 3. Independent Sample T-Test between Students Receiving Pull-out Resources and Push-In Resources..................................................................................................................33
Table 4. Independent Sample T-Test between students receiving Pull-out resources and Push-In Resources ..................................................................................................................35
Table 5. Independent Sample T-Test between Students Receiving Pull-out Resources and Push-In Resources..................................................................................................................36
Table 6. Independent Sample T-Test between students receiving Pull-out resources and Push-In Resources ..................................................................................................................38
Table 7. Independent Sample T-Test between Students Receiving Pull-out resources and Push-In Resources ..................................................................................................................39
Table 8. Independent Sample T-Test between Students Receiving Pull-out Resources and Push-In Resources ..................................................................................................................40
Table 9. Independent Sample T-Test between Students Receiving Pull-out Resources and Push-In Resources ..................................................................................................................42
Table 10. Independent Sample T-Test between Students Receiving Pull-out Resources and Push-In Resources ..................................................................................................................43
Table 11. Independent Sample T-Test between Students Receiving Pull-out Resources and Push-In Resources

Table 12. Independent Sample T-Test between Students Receiving Pull-out Resources and Push-In Resources

Table 13. Independent Sample T-Test between students receiving Pull-out resources and Push-In resources, GRADE 2-Reading portion of the WIDA

Table 14. Independent Sample T-Test between Students Receiving Pull-out Resources and Push-In Resources

Table 15. Independent Sample T-Test between Students Receiving Pull-out Resources and Push-In Resources

Table 16. Independent Sample T-Test between Students Receiving Pull-out Resources and Push-In Resources

Table 17. Independent Sample T-Test between Students Receiving Pull-out Resources and Push-In Resources Writing

Table 18. Independent Sample T-Test between Students Receiving Pull-out Resources and Push-In Resources

Table 19. Independent Sample T-Test between Students Receiving Pull-out Resources and Push-In Resources

Table 20. Independent Sample T-Test between Students Receiving Pull-out Resources and Push-In Resources
List of Figures

Figure 1. Kindergarten Listening ................................................................. .71
Figure 2. Kindergarten Speaking ................................................................. .72
Figure 3. Kindergarten Reading ................................................................. .73
Figure 4. Kindergarten Writing ................................................................. .74
Figure 5. Kindergarten Overall ................................................................. .75
Figure 6. Grade 1 Listening ................................................................. .76
Figure 7. Grade 1 Speaking ................................................................. .77
Figure 8. Grade 1 Reading ................................................................. .78
Figure 9. Grade 1 Writing ................................................................. .79
Figure 10. Grade 1 Overall ................................................................. .80
Figure 11. Grade 2 Listening ................................................................. .81
Figure 12. Grade 2 Speaking ................................................................. .82
Figure 13. Grade 2 Reading ................................................................. .83
Figure 14. Grade 2 Writing ................................................................. .84
Figure 15. Grade 2 Overall ................................................................. .85
Figure 16. Grade 3 Listening ................................................................. .86
Figure 17. Grade 3 Speaking ................................................................. .87
Figure 18. Grade 3 Reading ................................................................. .88
Figure 19. Grade 3 Writing ................................................................. .89
Figure 20. Grade 3 Overall ................................................................. .90
CHAPTER 1

INTRODUCTION

The United States of America is known as a “melting pot.” Within this melting pot, multiple languages are spoken, both at home and in the classroom. Meanwhile, more than 27 million immigrants have settled in the United States, causing classroom size to grow. Meanwhile, during this influx, for every 150 students, there is only one English as a Second Language (ESL) teacher. In other words, 10% of public school students are labeled as students who speak English as a Second language in comparison to only 1% of teachers who are trained to teach these students. In 2013, there were 63,739 English Language Learners (ELL) in New Jersey Public Schools, which is nearly 1 out of every 21 students.

According to the New Jersey Department of education, there are one hundred and sixty six different languages spoken by New Jersey students, and roughly 23% of these students are from a home where a language other than English is spoken. In New Jersey we have almost five out of six school districts have English Language Learners in their schools, within ten years the number grew over 30%. In addition, Passaic County had the largest population in 2013, at 12.47% or 10,023 in comparison Sussex County, which had the least, at 0.33%, or 73 students.

In fact, “During the 2007–08 school year, English learners represented 10.6% of the K–12 public school enrollment, or more than 5.3 million students. In fact, English learners are the fastest-growing segment of the student population, with their growth highest in grades seven through twelve” (Slavin, Calderón, & Sánchez, 2011). In addition to the two counties mentioned, the United States Census Bureau states that 30% of the population in New Jersey has someone ages of the ages of five or older living in a household speaks another language in addition to English.
The increasing population of ESL students has sparked educators as to which method is the best to teach these students. The continuation of high stakes testing places a huge amount of pressure on school districts in general. It is interesting to find out that fifteen states do not require that all teachers have expertise or training in working with ELLs, with New Jersey amongst those states, it creates a growing population of teachers that are unprepared to meet the needs of this growing population of students. To better understand what methods benefit the students the most, it is important to study the different types of methods when teaching English as a Second Language. Currently, the two main types of instruction are “Push-in” and “Pull-out.” The Pull-out instruction can create conflict where ESL specialists and mainstream teacher is complicated by hierarchical relationships of power, perceptions of status, and differences in the environment of instructional experience in the schools. In turn, consequently, Pull-out instruction possibly causes students to feel alienated and to miss class work, sometimes bringing about anxiety towards the unfinished work.

This study will compare and analyze the ESL program results according to the WIDA (World-Class Instruction Design and Assessment) ACCESS (Assessing Comprehension and Communication in English State-to-State for English Language Learners), which is an end of year examination used in two New Jersey school districts. Student performance is based on the methods of the program delivery. In addition, I will examine the different types of programs, which are mainly Push-in ESL: “Rather than pulling students out to teach them English, we would teach, them English in the mainstream” (Jewell, 2016). They would build their conversational and academic English-language skills in the general-education classrooms. There would be, in the new rhetoric of education, a "paradigm shift" in how we deliver instruction to limited-English-proficient students. Whatever is currently taught in the pullout class would now
also be taught in the regular classroom. This is possible to do, but, in recent years, some E.S.L. educators have proposed, instead, “Push-in” programs at the elementary level (Education Week, n.d.). We also have ESL Pull-out classes. “Pull-out English as a Second Language programs remove students from mainstream classrooms for a portion of the day in order to give them specialized instruction in English” (McKeon, 1987). “Pull-out ESL is most common in elementary schools where a designated ESL teacher works with small groups of children” (Rennie, 1993). Another aspect of this research involves understanding what criteria districts use to determine which program is best for their students.

Due to the small sample size of the groups, the results may not be generalizable beyond the specific population studied. Similarly, the time available to perform this study can only provide a snapshot of findings. Included in the delimitations are the groups to be sampled in both school districts. In addition, results that may occur in the two districts sampled may not have the same outcomes as districts in different parts of the country. Speculations may include the curriculum used, different teaching styles, student demographics, and parental involvement.

**Definition of Terms**

ACCESS for ELLs®. Assessing Comprehension and Communication in English State-to-State for English Language Learners. This large-scale test, first and foremost, addresses the English language development standards that form the core of the WIDA Consortium’s approach to instructing and testing English language learners. These standards incorporate a set of model performance indicators (PIs) that describe the expectations that educators have of ELL students at four different grade level clusters and in five different content areas (WIDA, n.d.).
The American Council on the Teaching of Foreign Languages (ACTFL). An organization dedicated to the improvement and expansion of the teaching and learning of all languages at all levels of instruction. ACTFL is an individual membership organization of more than 12,000 language educators and administrators from elementary through graduate education, as well as government and industry (American Council on the Teaching of Foreign Languages, n.d.).

Affective Filter. A "wall" that a learner puts up if his/her discomfort level is high. In ELLs, it is usually due to language proficiency and dealing with the process of adjusting to a new culture. ELLs need to have a low affective filter in order to learn English. The more comfortable students are in their school environment, the more ready they will be to learn (Linguistics Girl, n.d.).

Basic Interpersonal Communication Skills (BICS). The language skills needed for everyday personal and social communication. Second language learners must have BICS in order to interact in various types of settings. It usually takes students from one to three years to completely develop their social language. BICS are not necessarily related to academic success (Everything ESL, n.d.).

Beginning Sheltered Instruction (BSI). A course given only to students at ELP level 1 (New Jersey Department of Education, n.d.).

Cognitive Academic Language Proficiency (CALP). The language associated with native language literacy and cognitive development, which are the language skills needed to undertake academic tasks in the mainstream classroom. CALP includes content-specific vocabulary. It may take students from 5 to 7 years to develop CALP skills. CALP developed in
the first language contribute to the development of CALP in the second language (Everything ESL, n.d.).

**Comprehensible Input.** Input/instruction that is just above the students’ abilities. Instruction that is embedded in a meaningful context, adapted (e.g., paraphrasing, repetition), collaborative/ interactive and multimodal, and presented to students in a comprehensible way. Input + 1/Zone of Proximal Development (ELD Strategies, n.d.).

**Cognates.** Words in different languages that have common origin, related by descent from the same ancestral language by derivation, borrowing, or descent, usually related to a verb by derivation, and serving as its object to reinforce the meaning. Cognates are often pronounced in a similar way and have a meaning and spelling that are the same or close.

**ELL/ EL.** English Language Learners/English Learners (Mason, 2011)

**English Language Development (ELD).** A system of instruction focused on teaching ELLs to use English proficiently to communicate for various purposes in four language domains – speaking, listening, reading, and writing. ELD is also a class period to which that all the students placed in the ESL Program are assigned. It has its own curriculum and state standards (WIDA, n.d.).

**English Language Proficiency (ELP).** Levels of English language learners’ fluency, based on their stage of language acquisition and characterized by specific language behaviors in reading, writing, listening, and speaking. The levels are determined by the State ELPA Test. Level 1 is beginner, Level 2 is early-intermediate, Level 3 is intermediate, Level 4 is early-advanced, and level 5 is advanced (WIDA, n.d.).

**ESL.** English as a Second Language.
Exited. Refers to students who earned a composite score of 5 on ELPA are considered to be proficient in English and are exited from the ESL program; they will not receive an ELD class ("ACCESS for ELLs™," n.d.).

Five content areas. In the standards, the first is called social and instructional language (SI), which incorporates the proficiencies needed to deal with the general language of the classroom and the school. The others are English language arts (LA), math (MA), science (SC), and social studies (SS) (WIDA, n.d.).

Four language domains. Listening, speaking, reading, and writing (WIDA, n.d.).

Fluency. Automaticity and ease of comprehension (listening and reading) and production (speaking and writing); being able to use language for a variety of purposes (the same as proficiency) (Wisconsin Department of Public Instruction, n.d.).


Grade level clusters. These include Pre K-K, 1-2, 3-5, 6-8, and 9-12 (WIDA, n.d.).

Language Skills Transfer. The process of applying (transferring) language skills that students have learned in their first language (L1) into the second language (L2) (Hauptman, Mansur, & Tal, 2008).

Levene's Test. A test used to determine if k samples have equal variances. Equal variances across samples are called homogeneity of variance. Some statistical tests, such as the analysis of variance (ANOVA), assume that variances are equal across groups or samples. The Levene test can verify that assumption (1.3.5.10. Levene Test for Equality of Variances, n.d.).

Monitored. Students who have been promoted or exited from the ESL program are monitored for two additional years to ensure that language is no longer a barrier to their accessing content in an academic classroom (Colorín Colorado, n.d.).
Newcomers. Students who have recently arrived into the United States who have not been (or have only minimally been) exposed to communication in English. They usually require a specific intervention curriculum for a limited period. Newcomers might have had a varying degree of English as a Foreign Language instruction in their home countries. Newcomers may be exempt from state testing, depending on their PPS enrollment date (CLAS, n.d.).

Oral Proficiency Interview (OPI). A live, 20-30 minute conversation, taking place over the phone, between a trained, certified ACTFL tester and the candidate, the OPI is a valid and reliable test that measures how well a person speaks a language. The procedure is standardized in order to assess global speaking ability and measures language production holistically by determining patterns of strength and weakness. Through a series of personalized questions, a sample of speech is elicited and then rated against the proficiency levels described in ACTFL Proficiency Guidelines 2012. The primary goal of the OPI is the efficient elicitation of a ratable sample. To be ratable, a speech sample must clearly demonstrate the speaker’s highest sustained level of performance (known as the “floor”) and the level at which the speaker can no longer sustain the performance (known as the “ceiling”). The OPI resembles a conversation, but the tester respects a strict elicitation protocol and structures the interview in a specific way (Language Testing International, n.d.).

Proficiency Levels. Levels of English language learners’ proficiency (fluency), measured based on their stage of language acquisition and characterized by specific student language behaviors in reading, writing, listening, and speaking. The levels are determined by State ELPA Test. Level 1 is beginner, level 2 is early-intermediate, level 3 is intermediate, level 4 is early-advanced, and level 5 is advanced (WIDA, n.d.).
Promoted. ELLs who have earned a composite score of less than 5 on ELPA but demonstrated their language proficiency to be academically successful are promoted from the ESL program and do not receive an ELD more but are monitored for two years instead (Wisconsin Department of Public Instruction, n.d.).

Refusal of Services. Parents/guardians of the students eligible for ESL services may refuse such services being provided to their children. In this case, their child cannot enter an ELD classroom (New Jersey Department of Education, n.d.).

Sheltered Instruction (SI). An approach used to make grade-level academic content accessible to students with limited English proficiency while at the same time promoting their language development. Core curriculum is used but, in addition, language development functions and forms are also addressed. A carefully designed sequence of instructional activities is followed, and specific teaching strategies are implemented (What is Sheltered Instruction?, n.d.).

Statistical Package for the Social Science (SPSS). One of the most popular statistical packages that can perform highly complex data manipulation and analysis with simple instructions (Abbott, 2011).

Suspension of Services. It can be implemented after a student has been receiving ESL services for a number of years but still lacks language proficiency due to a learning disability. Suspension of services is a team decision and requires an informed consent of a parent or guardian (New Jersey Department of Education, n.d.).

W-APT. Stands for the WIDA-ACCESS Placement Test. This is an English language proficiency "screener" test given to incoming students. It assists educators with programmatic placement decisions, such as identification and placement of ELLs. The W-APT is one component of WIDA's comprehensive assessment system (WIDA, n.d.).
Writing Proficiency Test. (WPT) is a standardized test for the global assessment of functional writing ability in a language. The ACTFL writing proficiency tests measure how well a person spontaneously writes in a language (without access to revisions and/or editing tools) by comparing his/her performance of specific writing tasks with the criteria stated in the ACTFL Proficiency Guidelines – Writing. ACTFL writing tests are carefully constructed assessments with four to five requests for written responses dealing with practical, social, and/or professional topics encountered in informal and formal contexts. The writer selects between tasks and contexts that represent the range of proficiency levels on the ACTFL Proficiency Guidelines Writing – 2012. The tasks and prompts are in English with the expectation that the responses be written in the target language. ACTFL writing tests assess writing proficiency in terms of the ability to write effectively and appropriately for real-life writing purposes. They do not address when, where, why, or the way in which an individual learned to write. ACTFL writing tests are not an achievement test assessing a writer’s acquisition of specific aspects of course and curriculum content, nor are they tied to any specific method of instruction. ACTFL writing tests do not compare one individual’s writing to another individual’s writing, but rather compares each individual’s writing to the descriptors for writing (Language Testing International, n.d.).

WIDA. World-Class Instruction Design and Assessment (WIDA, n.d.).

In Chapter 2, the research will cover the different types of language programs, including the history and the types. In addition, it discusses the advantages and disadvantages of Push-in classes and classes where the students mainstream into the system and receive supplemental services in their classrooms.
CHAPTER 2
LITERATURE REVIEW

Chapter 2 provides the context for an overview of aspects of ESL/ELL, including the history of the program and the types of programs that exist. During the 1700s, “European immigrants settle in rural areas and open their own schools in their native languages and Founding US leaders decide against having a policy on language” (TimeRime.com, n.d.).

During the 19th Century, immigrants began flocking the United States in order to attain life, liberty, and the pursuit of happiness. In 1839, Ohio became the first state to adopt a bilingual education law, authorizing German-English instruction at parents' request. Louisiana enacted an identical provision for French and English in 1847, and the New Mexico Territory did so for Spanish and English in 1850. By the end of the 19th century, about a dozen states had passed similar laws. Elsewhere, many localities provided bilingual instruction without state sanction, in languages as diverse as Norwegian, Italian, Polish, Czech, and Cherokee. (Rethinking Schools Online, n.d.)

This influx caused the United States to pass the “Naturalization Act in 1906, which was instrumental in requiring the knowledge of English in America. The Act required that all immigrants be able “to speak English in order to become naturalized citizens of the U.S” (The History of ESL, n.d.). Later, “by the mid-1920s, bilingual schooling was largely dismantled throughout the country. English-only instruction continued as the norm for LEP students until its failure became evident. LEP students in English-only classrooms were falling behind in their academic studies and dropping out of school at alarming rates” (Rethinking Schools Online, n.d.).
A huge step for equality in bilingual education was the Civil Rights Act of 1964. This act forbade types of discrimination based on sex as well as race, color, or national origin in hiring, promoting, and firing. The passing of the Civil Rights Act paved the way for the Elementary and Secondary Education Act of 1965 (ESEA), which was a Great Society program enacted in 1965 that allocates federal funding for primary and secondary school education and forbs the establishment of a national curriculum. The Elementary and Secondary Education Act also provided a mechanism to hold schools accountable and increase equality in education nationally.

The current reauthorization of this bill is the No Child Left Behind Act, signed into law in 2001 (Elementary And Secondary Education Act, n.d.). Included under this act are the children of low-income family, homeless, neglected, migrant, Native American, and English language limited families. The “Secondary Education Act aims to provide for their long-term welfare by improving their schools and the resources available to them. In 1965, when this Act became law, there was a large “achievement gap” stratified by race and poverty” (Elementary and Secondary Education Act, n.d.).

Another provision of this act was Title Seven, which provided federal funding to bilingual education and specialized programs for the advancement of Native Hawaiians, Alaskans, and Native Americans. The most important legacy of the Bilingual Education Act, introduced under the Elementary and Secondary Education Act, was the establishment of separate bilingual classes for English Language Limited students, acknowledging that educating them alongside proficient students was detrimental to their long-term development. The bilingual educational program prepared these students to join their English-speaking peers as soon as possible. Not a separate form of education, it is a remedial practice to prepare students for the
regular educational system. This also meant that putting students in these classes was not “segregation,” which had been outlawed a few years earlier.

Subsequent amendments to the Bilingual Education Act outlined specific goals for programs and clarified the need for these programs to overcome language barriers. In the current revision of the Bilingual Education Act, this title was incorporated as Title III Language Instruction for Limited English Proficient and Immigrant Students (Elementary and Secondary Education Act, n.d.). Soon after, the Bilingual Education Act of 1968 created another means of aiding those students who speak another language to learn English.

The need for bilingual education grew as non-Native English speaking students began to flood the public school system. In 1978, the Bilingual Education Act funded more than 518 bilingual and bicultural projects in sixty-eight different languages, including Spanish, Chinese, and Polish. By 1983, the National Association for Bilingual Education was pushing students to receive bilingual education and not to maintain their primary language. This was a way for the National Association for Bilingual Education Act to make the primary language English. Many teachers supported the National Association for Bilingual Education Act because they felt that the maintenance of their primary language impeded the students’ English acquisition.

In the 21st century, No Child Left Behind Act of 2001 (NCLB), which was the reauthorization of the Elementary and Secondary Education Act of 1965, distributes funds to states to improve the education of limited English proficient students by supporting these children to learn English and meet rigorous state academic content and achievement standards. Title III of No Child Left Behind Legislation provides support for limited English proficient students.

Other ESL program types include Sheltered English Instruction, which is an instructional
approach used to make academic instruction in English understandable to LEP students. Regular classroom teachers teach Sheltered English classes and receive training on strategies to make subject area content comprehensible for LEP students (State of New Jersey Department of Education, n.d.)

A full-time bilingual program with ESL Support is required for instruction in all subjects that a child is required by law or statute to master, to be administered in the native language of the limited English proficient student and in English. A bilingual program is required when there are 20 or more limited English proficient students in any one-language classification enrolled in the school district. An example of this would be a school district with more than 200 Spanish-speaking students and more than 20 bilingual-certified teachers that offers a self-contained program for grades K-8 and a departmentalized program for grades 9-12. The use of native language for instruction is influenced by the English language proficiency and academic level of the student in their native language (State of New Jersey Department of Education, n.d.)

A part-time bilingual program includes students who are mainstreamed into English program classes but receive daily instruction in mathematics and reading by a certified bilingual teacher. This occurs, for example, when a school district with approximately 30 bilingual Spanish students, 30 bilingual Portuguese students, and 2 bilingual-certified teachers offers a part-time pullout bilingual resource room. Students receive a minimum of one period of developmental reading instruction and support in math from the bilingual teacher, in addition to their regular ESL instruction (State of New Jersey Department of Education, n.d.)

In a high-intensity ESL program, students receive two or more periods of ESL instruction per day. One period is the standard ESL class and the other period is tutorial or ESL reading class. For example, a school district with approximately 25 bilingual Spanish students and 3 ESL
teachers might offer high-intensity ESL. Students receive two class periods of ESL instruction. The teachers provide push-in instruction for content area classes and daily Pull-out instruction for ESL (State of New Jersey Department of Education, n.d.)

In an ESL-only program, a daily developmental second-language program of up to two periods of instruction based on student’s needs that school districts must provide when there are 10 or more LEP students enrolled in the district (North Jersey Arts & Science Charter Schools, n.d.).

English Language Services (ELS) programs are provided in districts with fewer than 10 students of limited English proficiency in order to improve their English language skills. For example, school districts with eight ELS students are provided approximately 90 to 100 minutes of ELS instruction weekly. Students can integrate with the student body, but they also receive ELS instruction weekly from a certified teacher in the district (State of New Jersey Department of Education, n.d.).

The ESL Pull-out program under No Child Left Behind required that all public schools help ESL students to become English proficient, as both fluent and ESL students are mandated to meet state and national achievement standards. While public schools are legally required to provide educational support for ESL kids, many experts assert that the current approach is both ineffective and flawed. As the number of minority students is rising each year in the United States, experts predict that by the year 2020, at least 50% of school-age children will be of non-English speaking backgrounds. With the current and rising statistic of non-English speaking students in public schools, leaders assert that all schools must take a more active approach to protecting the instructional support for these non-English speaking students (PublicSchoolReview.com, n.d.).
The state mandate requires that students receive language classes, but it does not state which programs are required or which are beneficial. Pull-out ESL classes provide students with the opportunity to work is small group instruction outside of the classroom, however:

One of the biggest advantages of Pull-out ESL classes is that the groups are usually much smaller than the grade-level classroom. When students receive instruction in small groups, they naturally also receive more of the teacher's attention. In the ESL Pull-out classroom, this means the students have more opportunities to practice speaking when giving answers, as there are fewer children competing to talk. Likewise, small classes mean reluctant students cannot "fly under the radar," or quietly not participate. Students are also less likely to become distracted since there is less activity than in a large classroom” (Everyday Life - Global Post, n.d.).

Here, students work on skills that they will need in a designated area in order to acquire the language more easily.

Pull-out also allows for targeted instruction, which, in turn, provides the opportunity for students to be leveled at their language level. Working together at the same language level gives the students the opportunity to move at their own pace and work together as a group. “This means the ESL teacher can work with low-level students on the basics and high-level students on grade-level language. Thus, the teacher provides more examples of the target structures and modifies speech appropriate to the students. Either way, the students receive intense, individual attention focused on increasing their language skills (Everyday Life - Global Post, n.d.).

Students working within this small group help each other grow into a community. This community provides students with a safe and non-judgmental environment.
ESL Pull-out classes provide a safe space for non-native English speakers to practice their linguistic skills. Indeed, ESL students often report feeling frustrated in the grade-level classroom, not understanding everything and, therefore, unable to participate fully. Conversely, in the ESL class, the teacher modifies the language so that it is appropriate to the students' levels, thus providing them with opportunities to participate. Since the students feel comfortable, they are more likely to take educational risks, thus speeding up their language acquisition. Students in ESL Pull-out classes also report feeling a sense of community that comes from being grouped with others who can understand their specialized needs” (Everyday Life - Global Post, n.d.).

Not all educators support Pull-out, however. Some researchers have found that students who are pulled out of their classes for this extra support can feel left out. At times, these students can also feel punished because of the segregation, and/or that it may cause other students to look at them differently. According to Education Week, while support programs are, in many ways, essential for ESL children and families, many believe that separate setting instructional methods, which involve separating ESL children from fully fluent classrooms for the entire or partial class period, inappropriately create an environment of segregation.

Push-in ESL allows the students to remain in the classroom and to follow along with the rest of the class while getting extra support. “The rationale for Push-in has never rested on research findings, but on principle,” however. “Proponents insist that the integration of students with disabilities is inherently right, compared often to the same right to racial integration. The generally accepted concept of Push-in is that students with disabilities attend classes with their general education peers with direct support from special educators” (Colorin Colorado, n.d.). According to Kochhar, West, and Taymans (2000), that the benefits of Push-in across grade
levels far outweigh the difficulties that Push-in presents. For example, they believe that for students with disabilities, Push-in does the following:

1. facilitates more appropriate social behavior because of higher expectations in the general education classroom;
2. promotes levels of achievement higher or at least as high as those achieved in self-contained classrooms;
3. offers a wide circle of support, including social support from classmates without disabilities; and
4. improves the ability of students and teachers to adapt to different teaching and learning styles.

One argument is that, in some cases, the teachers must rethink their design of teaching and, sometimes, slow down the pace that they are using because students who do not speak English learn at a slower pace, providing the rest of the class with the opportunity to catch up with what the teacher is doing.

In addition to the argument that specialized support helps to increase ESL student learning, many educators and parents argue that the additional ESL support helps to maintain the challenging environment of a fluent classroom. Specifically, as a regular paced English class is designed to instruct students who are of an “average” English speaking proficiency, immersing non-English speaking students in such classes often forces teachers to slow down the pace of instruction, making the overall challenge of the class less rigorous for English speaking students.” (PublicSchoolReview.com, n.d.) However, some argue that this less rigorous environment does not benefit students whose abilities are higher than others’; here, the students can become uninterested. In addition, some
believe that Push-in does not benefit any of the students in the class. In the article “Success with ELLs,” Sarah Cacicio focuses on a student who came from the Dominican Republic and who, despite having three years of English instruction in the Dominican, was still having problems understanding, reading, and communicating in English. Next, teachers decided to use a different method than a Push-in/pull out method: the Collaborative Systems for Learning Model, a flexible model for collaborative student research in coursework across the curriculum. The foundation of the model lies in its cooperative stance, which supports students working together toward informed decision making about a common research problem. Its use started with communication: short conversations in passing, which led to meetings.

Brief meetings then led to collaborating on lesson plans, assignments, language objectives, and more appropriate methods of assessment. For example, the ESL co-teacher and I met with Pablo’s math teacher, who ordered the same-level curriculum textbook and additional materials in Spanish. They reviewed the student’s progress and agreed to let him to take the standardized math tests in his home language. Allowing Pablo to use both languages in school also gave value to Spanish as an academic language, encouraging greater equality in terms of culture and curriculum. Pablo began to show improvement at the same rate as his English-speaking peers.

Pull-out and Push-in ESL each have supporters and those who oppose them. Research has yet to prove which is more beneficial to students. Students, in both methods, can feel excluded in the classroom. On the other hand, both can benefit the students with the extra support.

Chapter 3 identifies the type of research design that is to be used in this study. The design will include the WIDA (World-Class Instructional Design and Assessment) in listening, reading, writing, and speaking scores of kindergarten through 3rd grade ELL students who participate in a
Push-in model of ESL compared against those K-3 students who participated in a Pull-out type of class. Two adjoining school districts in Bergen County, New Jersey were used to generate the findings.

In addition, due to the lack of research on this specific topic, articles and findings were not easily accessible. Dissertations and peer review articles were limited. Although studies supported evidence in finding for both supporting and opposing both methods, results were minimal. Further research broadening the topic including larger sample sets could possibly expand research on the topic of ESL teaching methods.
CHAPTER 3

METHODOLOGY

This research design for this study was a causal-comparative design to evaluate the effects of the WIDA (World-Class Instructional Design and Assessment) on the listening, reading, writing, and speaking scores of kindergarten through third grade ELL students who participate in an Push-in model of ESL compared to those students, K-3, who participate in an Pull-out type of class. This research design will analyze the results of two school district's ESL programs according to the WIDA (World-Class Instruction Design and Assessment) ACCESS (Assessing Comprehension and Communication in English State-to-State for English Language Learners) end-of-year examination. Student performance was measured based on the methods of the program delivery: either “push-in” or “Pull-out” services.

In addition, I examined the different types of programs, mainly Push-in ESL, in which the students learn English in the mainstream classroom and build their conversational and academic English-language skills in the general-education classroom. “There would be, in the new rhetoric of education, a ‘paradigm shift’ in how we deliver instruction to limited-English-proficient students. Whatever is now taught in the Pull-out class would be taught in the regular classroom. This is possible to do, but, in recent years, some E.S.L. educators have proposed instead “pull in’ programs at the elementary level” (Education Week, n.d.). We also have ESL Pull-out classes, which “remove students from mainstream classrooms for a portion of the day in order to give them specialized instruction in English” (McKeon, 1987). “Pull-out ESL is most common in elementary schools where a designated ESL teacher works with small groups of children” (Rennie, 1993). Another aspect that I examined is what criteria districts use to determine which programs are best for their students.
A causal-comparative design method is best suited for this study because

a causal-comparative design is a research design that seeks to find relationships between independent and dependent variables after an action or event has already occurred. The researcher's goal is to determine whether the independent variable affected the outcome, or dependent variable, by comparing two or more groups of individuals. There are similarities and differences between causal-comparative research, also referred to as ex post facto research, and both correlational and experimental research. This section discusses these differences, as well as the benefits, process, limitations, and criticism of this type of research design. To demonstrate how to use causal-comparative research, examples in education are presented. Many similarities exist between causal-comparative research and correlational research. Both methods are useful when experimental research has been deemed impossible or unethical as the research design for a particular question.” (Salkind, 2010)

More specifically, I used a t-test of Independent Samples, in which groups are chosen independently such that choosing the subjects for one group has nothing to do with choosing the subjects for another group. Additionally, the t-test of Independent samples accounts for the equality of variances, or differences. Specifically, the Levene’s test is used to test for these differences. We are to assume that variances from the populations are equal, in this instance the Levene’s test tests this assumption specifically that the null hypothesis that the population of the variances are equal. If the p-value is less than significant, in this case .05, the obtained in the differences in samples is unlikely to have occurred because we have random sampling from a population with equal variances. Therefore, the null hypothesis of equal variances is rejected and it is stated that there is a difference between the variances in the samples. Finally, I compared
the post-test between the two treatment groups. “The Chief concern with this test is the difference in the means of the two samples” (Abbott, 2011).

My reasoning for this approach is as follows:

1. I have two groups of students, some who learn English by being pulled out of the class and some who learn by teachers pushing them into the class;
2. I have to assume that the students were similar before they started receiving English services;
3. My task is to determine which style of teaching affects the students’ test scores.
4. If the scores are significantly different, I can say that the teaching style does affect the test scores directly.

The archival data used were the scores of the ELL students on the reading, listening, speaking, and writing sections of the WIDA ACCESS. The districts administer the assessments annually between the months of March and May; the specific testing window is according to state mandates.

**Setting and Sample**

The criteria used for this research was students in kindergarten through 3rd grade over three years, comparing two school districts. A sample set of 30 to 60 students from each grade level was chosen at random. Each school district will be using one of these methods for each of the grades, logging back three years. District A and B (the two districts whose test scores will be compared) each has a population of students who speak another language at home; both districts A and B have over 40% of their students who speak another language at home.

According to Menken and Kleyn (2010), “Researchers into second language acquisition have estimated that it takes between 5 to 7 years for an ESL student to catch up with his native-
speaking peers as far as academic English is concerned”. Using the Push-in and Pull-out method allows each district to accommodate their students and provide the needed type of services and differentiation to meet their students’ needs.

District A is slightly less than one square mile in area and is located eight miles from New York City on the banks of the Passaic River. The district has three schools serving 1,250 students: School (K-3), School (K-6), and Jr./Sr. High School (7-12). District A reports the percentage of students who primarily speak each language in their home as follows: English 40.0%, Polish 37.7%, Spanish 9.3%, Arabic 4.4%, Gujarati 2.7%, Macedonian, 1.7%, and Other 4.4%. As of the 2011-12 school years, District B's 11 schools had an enrollment of 4,500 students. Students who primarily speak each language at home are English 48.9%, Spanish 30.3%, Polish 9.3%, Albanian 3.2%, Macedonian 3.1%, Arabic 1.2%, and other 4.2%. These reports show that there is a large school population that does not speak English as a primary language at home.

**Instrumentation and Resources**

The WIDA ACCESS is a state mandated assessment that servers as an additional requirement for students who are enrolled in ELL classes to be exited from the program. In addition, this assessment follows the students and allows educators to track progress in the four domains of listening, speaking, reading, and writing:

ACCESS for ELLs (Assessing Comprehension and Communication in English State-to-State for English Language Learners) is a secure large-scale English language proficiency assessment given to students in Kindergarten through 12th graders who have been identified as English language learners (ELLs). The WIDA ACCESS is given annually in
WIDA Consortium member states to monitor students' progress in acquiring academic English” (WIDA, n.d.).

Children are administered the WIDA ACCESS placement test for a particular grade level cluster one semester beyond the beginning of the grade level cluster and one semester beyond the end of the grade level cluster. For example, the grades 3-5 W-APT (WIDA-ACCESS placement test) is administered to children from 2\textsuperscript{nd} semester grade 3 through 1\textsuperscript{st} semester grade 6. The W-APT (WIDA-ACCESS placement test) for all grade levels except kindergarten is administered as follows:

1. The Speaking test is administered first, and then followed by the Listening,
   Reading, and Writing tests.

2. All components are adaptive, meaning that the component in process is stopped when the test administrator reaches items with more complex material than the student is likely to successfully handle.

Instructions for stopping a component are included in the “test script” (WIDA, n.d.). For example, if a student does not answer a sufficient number of questions correctly in a certain section, then the section is completed and the student moves on to the next portion of the test.

The reports can be used along with other information as a source of data to help understand systems-level strengths and needs. In addition, according the ACCESS for ELLs, Growth Reports are intended to help educators identify district- and school-level patterns in language growth (WIDA, n.d.). These reports are available to districts and schools at an additional cost and they include a) growth data for students with two years of test scores for multiple two-year-spans, b) growth data for every domain and composite at each grade level, c) district-level and school-level reports, d) charts showing students’ growth compared to WIDA-
wide growth, e) tables that show the number and percent of students with two years of scores, and f) downloadable CSV files of student data. The growth reports should not be used for federal, district, or school accountability purposes, teacher evaluation, or individual student progress monitoring.

Districts identify the students who will receive the ESL instruction in September or the beginning of the school year. After parental consent, students are placed in NJ SMART (New Jersey Standards Measurement and Resource for Teaching) and are identified in the district as receiving services. In January, districts send in the level of test they feel that each student should receive. In kindergarten, all the levels are the same; in grades 1 through 12, the levels are A, B, and C, with A being the lowest level and C the hardest. Additionally, districts will receive the test materials within the same month. Districts are asked to test within the March through April window and submit final tests in April. Reports of the test should reach the districts by the end of June.

Students who meet the criteria are exited out according to grade level. For grades 1 through 12,

The regulation mandating the use of multiple criteria for identifying and exiting students from language assistance programs services is still current. For students in grades kindergarten through 12, a W-APT (WIDA-ACCESS placement test) proficiency level of 4.5 is the standard for determining the need for language assistance services. Students who score below this standard and have at least one other indicator, are students identified as limited English proficiency (ACCESS for ELLs™, Entry and Exit Criteria on W-APT Kindergarten Test, n.d.).
In addition, educators are urged to use multiple measures when exiting a child from the program. WIDA ACCESS is not the only measure to determine if a student is eligible for services.

**Data Collection**

After the Institutional Review Board approved this study, data collection began. Score sheets containing the data were collected from grades K-3 backlogging three years to track progress on some students. Each score reported was blind, only showing the results of the four domains and the overall score. Typically, score results are reported yearly to district supervisors in order to send reports to the New Jersey Department of Education. The $t$-test of statistical significance was conducted once data were collected from both school districts. The $t$-test of Statistical Significance “indicates whether or not the difference between two groups’ averages most likely reflects a ‘real’ difference in the population from which the groups were sampled” ("T-Test Statistical Significance Example and Definition,"). The statistical method showed the significance of the differences. For this study, the independent variable was the Push-in and the Pull-out model used in both school districts. Scores collected using the four domains, and the overall total score, were the dependent variable in this study. Using this information, this study’s purpose was to determine whether there is a significant difference or positive results showing higher growth percentages in a district, dependent on the type of procedure.

The research question that guiding this study was as follows: Do students who participate in ESL Push-in classes benefit more than students who receive Pull-out ESL classes based on their WIDA ACCESS scores?

**Null Hypothesis Alternative Hypothesis**
**Null Hypothesis 1.** There is no significant difference between the performance scores on the listening portion of the WIDA ACCESS for ELLs who participate in a Push-in or Pull-out instructional program.

**Alternative Hypothesis 1.** There is a significant difference between the performances of scores on the listening portion of the WIDA ACCESS for ELLs who participate in a Push-in or Pull-out instructional program.

**Null Hypothesis 2.** There is no significant difference between the performances of scores on the reading portion of the WIDA ACCESS for ELLs who participate in a Push-in or Pull-out instructional program.

**Alternative Hypothesis 2.** There is a significant difference between the performances of scores on the reading portion of the WIDA ACCESS for ELLs who participate in a Push-in or Pull-out instructional program.

**Null Hypothesis 3.** There is no significant difference between the performances of scores on the speaking portion of the WIDA ACCESS for ELLs who participate in a Push-in or Pull-out instructional program.

**Alternative Hypothesis 3.** There is a significant difference between the performances of scores on the speaking portion of the WIDA ACCESS for ELLs who participate in a Push-in or Pull-out instructional program.

**Null Hypothesis 4.** There is no significant difference between the performances of scores on the writing portion of the WIDA ACCESS for ELLs who participate in a Push-in or Pull-out instructional program.
**Alternative Hypothesis 4.** There is a significant difference between the performances of scores on the writing portion of the WIDA ACCESS for ELLs who participate in a Push-in or Pull-out instructional program.

**Null Hypothesis 5.** There is no significant difference between the performance scores on the overall score of the WIDA ACCESS for ELLs who participate in a Push-in or Pull-out instructional program.

**Alternative Hypothesis 5.** There is a significant difference between the performances of scores on the overall score portion of the WIDA ACCESS for ELLs who participate in a Push-in or Pull-out instructional program.

**Assumption, Limitations, Delimitations**

An assumption behind this study was that all the teachers were certified and highly qualified. According to the NJDOE, “New Jersey requires that candidates for certification achieve a cumulative GPA of at least 3.0.” This equals an A grade for students graduating on or after September 1, 2016 (2.75 for those graduating before September 1, 2016) in a baccalaureate degree program, higher degree program, or a state-approved post-baccalaureate certification program with a minimum of 13 semester-hour credits. A high praxis test score may offset a GPA that is lower than 3.0 but higher than 2.75. Meanwhile, current regulations for certification require that applicants complete a state-approved English as a Second Language program culminating in student teaching. One must also present evidence of passing oral and written English language proficiency tests (OPI & WPT). These tests must be passed with a score of “Advanced Low” or higher (New Jersey Department of Education, n.d.).
Another expectation was that all the teachers administering each part of the test had been trained properly on how to administer the exam. Other factors to consider were the experience of each teacher, the level of education held, and whether there was any continuing education taking place in the field of ESL. The study was delimitated to $x$ amount of students in grades kindergarten through third grade in districts A and B.
CHAPTER 4

RESULTS

Demographics

As noted in Chapter 3, the make-up of the districts was as follows: District A is slightly less than one square mile in area, located eight miles from New York City on the banks of the Passaic River. The district has three schools serving 1,250 students: School (K-3), School (K-6), and Jr./Sr. High School (7-12). District A reports the percentage of students who primarily speak each language in their home as English 40.0%, Polish 37.7%, Spanish 9.3%, Arabic 4.4%, Gujarati 2.7%, Macedonian, 1.7%, and Other 4.4%. As of the 2011-12 school year, District B's 11 schools had an enrollment of 4,500 students. Those students who primarily speak each language at home were English 48.9%, Spanish 30.3%, Polish 9.3%, Albanian 3.2%, Macedonian 3.1%, Arabic 1.2%, and other 4.2%. The table below shows the number of students in each group. Race, gender, and other identifiers were not included because the samples were chosen at random to find only the effects of Push-in and Pull-out outcomes. The purpose of this study was to determine the most effective ESL intervention model comparing it to the effect of the “Push-in” or Push-in and the “Pull-out” or Pull-outs of students in the four domains: listening, speaking, reading, writing, and overall in grades kindergarten to three. The quantitative approach helped to determine if there was a statistically significant difference between listening, speaking, reading, writing, and overall score in WIDA results when comparing the models. Interpreting the final results using the independent sample $t$-test, we first must understand the requirements. The requirements for the independent sample $t$-test are a) the samples are independent of one another, and b) the dependent variable is interval level, sample populations
are normally distributed, and both populations have equal variance (this is also known as the “test of homogeneity” because we are assessing “sameness”) (Abbott, 2011).

SPSS provides a procedure for testing the equality of variances called the “Levene's Test.” Levene's test (Levene, 1960) is used to test if $k$ samples have equal variances. Homogeneity of variance refers to equal variance across samples. Some statistical tests, such as the analysis of variance, assume that variances are equal across groups or samples. The Levene test can be used to verify that assumption (1.3.5.10. Levene Test for Equality of Variances, n.d.). A description of the results follows.

**Kindergarten**

This sample set consists of 60 kindergarten students from both districts A and B ages 4 through 6. The students were chosen at random and their identities concealed, and the independent $t$-test was used to compare the scores from both districts. It was not clear what the gender, race, or language spoken at home was, and this information was not taken into consideration for the random data selection. Results of listening, speaking, reading, writing, and overall scores are as shown in Table 1.

**Null Hypothesis 1.** There is no significant difference between the scores on the listening portion of the WIDA ACCESS for ELLs who participate in a Push-in or Pull-out instructional program.

**Alternative Hypothesis 1.** There is a significant difference between the performances of scores on the listening portion of the WIDA ACCESS for ELLs who participate in a Push-in or Pull-out instructional program.
To reveal differences in performance scores on the WIDA ACCESS for ELLs who participate in a Push-in or Pull-out instructional program, an independent samples \( t \)-test was conducted. If Levene’s test for equality of variances is significant, the statistics for the row equal variances are not assumed, and the altered degrees of freedom are rounded to the nearest whole number. The level of significance was .05 with \( F(1,101.2)=17.5, p=.000 \) a \( t \)-test, not assuming calculated homogeneous variances. The results of this test indicated that there was a significant difference in weight observed between the two groups, \( t(101.2)=-4.49, p=.000 \), Pull-out \( (M=280.9; SD=76.5) \) and Push-In \( (M=333.8; SD=49.6) \), as shown in Figure 1. Therefore, I could reject the null hypothesis claiming there is no difference in test scores and accept the alternative hypothesis that there is a difference in scores between students in Push-in and Pull-out classes. As shown as in Figure 1, students given Push-in services received a higher score of about 50
points. This outcome could be due to the fact that, during Push-in services, students are in the classroom environment being exposed to other students, which allows them to listen to a larger number of children in their age range.

Table 2

*Independent Sample T-Test between students receiving Pull-out resources and Push-In resources.*

<table>
<thead>
<tr>
<th>Kindergarten-Reading portion of the WIDA</th>
</tr>
</thead>
<tbody>
<tr>
<td>N= Number of students</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Pull-out Resources</td>
</tr>
<tr>
<td>Pull-in Resources</td>
</tr>
</tbody>
</table>

**TABLE 2**

<table>
<thead>
<tr>
<th>Levene’s Test for Equality of Variances</th>
<th>t-test for Equality of Means</th>
</tr>
</thead>
<tbody>
<tr>
<td>Change in Equal variances assumed weight post-test</td>
<td>.005</td>
</tr>
<tr>
<td>Equal variances not assumed</td>
<td>.231</td>
</tr>
</tbody>
</table>

(Reading-Kindergarten)

Null Hypothesis 2. There is no significant difference between the scores on the reading portion of the WIDA ACCESS for ELLS who participate in a Push-in or Pull-out instructional program.

Alternative Hypothesis 2. There is a significant difference between the performances of scores on the reading portion of the WIDA ACCESS for ELLS who participate in a Push-in or Pull-out instructional program.

In order to examine differences in performance scores on the WIDA ACCESS for ELLs who participate in a Push-in or Pull-out instructional program an independent samples t-test was
conducted. If Levene’s test for equality of variances is significant, the statistics for the row equal variances are not assumed, with the altered degrees of freedom rounded to the nearest whole number. The level of significance was .05 with $F(1,118)=.005$, $p=.942$ a $t$-test assuming homogeneous variances was calculated. The results of this test indicated that there was not a significant difference in weight observed between the two groups, $t(.231)=118$, $p=.818$, Pull-out ($M=215.4; SD=63.7$) and Push-In ($M=212.7; SD=45.3$), as shown in Figure 2. Therefore, I could accept the null hypothesis claiming there is not a significant difference in test scores. In the Kindergarten Reading portion of the test, students’ scores are not dependent on the type of instruction they receive.

Table 3

*Independent Sample T-Test between Students Receiving Pull-out Resources and Push-In Resources*

<table>
<thead>
<tr>
<th>Kindergarten-SPEAKING portion of the WIDA</th>
<th>N= Number of students</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
</tr>
<tr>
<td>Pull-out Resources</td>
<td>60</td>
</tr>
<tr>
<td>Push-in Resources</td>
<td>60</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>TABLE 3</th>
<th>F</th>
<th>Sig.</th>
<th>t</th>
<th>df</th>
<th>Sig.(2-tailed)</th>
<th>Mean Difference</th>
<th>Std. Error Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Change in weight post-test</td>
<td>18.78</td>
<td>.000</td>
<td>-3.24</td>
<td>118</td>
<td>.002</td>
<td>-36.25</td>
<td>11.18</td>
</tr>
<tr>
<td>Equal variances assumed</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Change in weight post-test</td>
<td>-3.24</td>
<td></td>
<td>90.2</td>
<td></td>
<td>.002</td>
<td>-36.25</td>
<td>11.18</td>
</tr>
<tr>
<td>Equal variances not assumed</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(Speaking-Kindergarten)
Null Hypothesis 3. There is no significant difference between the performances of scores on the speaking portion of the WIDA ACCESS for ELLs who participate in a Push-in or Pull-out instructional program.

Alternative Hypothesis 3. There is a significant difference between the performances of scores on the speaking portion of the WIDA ACCESS for ELLs who participate in a Push-in or Pull-out instructional program.

In order to examine differences in performance scores on the WIDA ACCESS for ELLs who participate in a Push-in or Pull-out instructional program an independent samples t-test was conducted. If Levene’s test for equality of variances is significant, report the statistics for the row equal variances not assumed with the altered degrees of freedom rounded to the nearest whole number. The level of significance was .05 with $F(1,118)=18.78, p=.000$ a t-test not assuming homogeneous variances was calculated. The results of this test indicated that there was a significant difference in weight observed between the two groups, $t(90.2)=3.24, p=.002$, Pull-out ($M=309.1; SD=76.4$) and Push-In ($M=345.4; SD=40.1$). Therefore, I rejected the null hypothesis claiming there is no significant difference in test scores. In the Kindergarten Speaking portion of the test, students’ scores are dependent on the type of instruction they receive. As shown in Figure 3, students who were taught using a Push-in type of program received a higher mean score on the speaking portion of the WIDA ACCESS test.

Null Hypothesis 4. There is no significant difference between the performances of scores on the writing portion of the WIDA ACCESS for ELLs who participate in a Push-in or Pull-out instructional program.
**Alternative Hypothesis 4.** There is a significant difference between the performances of scores on the writing portion of the WIDA ACCESS for ELLs who participate in a Push-in or Pull-out instructional program.

Table 4

*Independent Sample T-Test between students receiving Pull-out resources and Push-In Resources*

<table>
<thead>
<tr>
<th>Kindergarten-WRITING portion of the WIDA</th>
<th>N= Number of students</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pull-out Resources</td>
<td>60</td>
</tr>
<tr>
<td>Mean</td>
<td>227.7</td>
</tr>
<tr>
<td>Std. Deviation</td>
<td>73.3</td>
</tr>
<tr>
<td>Std. Error Mean</td>
<td>9.46</td>
</tr>
<tr>
<td>Push-in Resources</td>
<td>60</td>
</tr>
<tr>
<td>Mean</td>
<td>242.0</td>
</tr>
<tr>
<td>Std. Deviation</td>
<td>48.7</td>
</tr>
<tr>
<td>Std. Error Mean</td>
<td>8.87</td>
</tr>
</tbody>
</table>

**Independent Samples Test**

<table>
<thead>
<tr>
<th>Levene’s Test for Equality of Variances</th>
<th>t-test for Equality of Means</th>
</tr>
</thead>
<tbody>
<tr>
<td>TABLE 4</td>
<td>F</td>
</tr>
<tr>
<td>Change in weight</td>
<td>.377</td>
</tr>
<tr>
<td>post-test Equal variances not assumed</td>
<td></td>
</tr>
</tbody>
</table>

(Writing-Kindergarten)

In order to examine differences in performance scores on the WIDA ACCESS for ELLs who participate in a Push-in or Pull-out instructional program, an independent samples *t*-test was conducted. If Levene’s test for equality of variances is significant, report the statistics for the row equal variances not assumed, with the altered degrees of freedom rounded to the nearest whole number. The level of significance was .05 with $F(1,118)=.377$ $p=.540$ a *t*-test assuming homogeneous variances were calculated. The results of this test indicated that there were not significant differences in weight observed between the two groups, $t(-1.10)=118$, $p=.272$, Pull-out ($M=227.7; SD=73.3$) and Push-In ($M=242.0; SD=48.7$) as shown in Figure 4. Therefore, I
accepted the null hypothesis claiming there is not a significant difference in test scores. In the Kindergarten Speaking portion of the test, students’ scores were not dependent on the type of instruction they receive.

Table 5

*Independent Sample T-Test between Students Receiving Pull-out Resources and Push-In Resources*

<table>
<thead>
<tr>
<th>Kindergarten-OVERALL portion of the WIDA</th>
<th>N</th>
<th>Mean</th>
<th>Std. Deviation</th>
<th>Std. Error Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pull-out Resources</td>
<td>60</td>
<td>243.8</td>
<td>61.5</td>
<td>7.93</td>
</tr>
<tr>
<td>Push-in Resources</td>
<td>60</td>
<td>261.3</td>
<td>53.7</td>
<td>6.93</td>
</tr>
</tbody>
</table>

---

**Null Hypothesis 5.** There is no significant difference between the performances of scores on the overall score of the WIDA ACCESS for ELLs who participate in a Push-in or Pull-out instructional program.

**Alternative Hypothesis 5.** There is a difference between the performances of scores on the overall score portion of the WIDA ACCESS for ELLs who participate in a Push-in or Pull-out instructional program.

The final part of the kindergarten test observed the overall score on the WIDA test. Here, the overall score consisted of 35% reading, 35% writing, 15% listening, and 15% speaking. In
order to examine differences in performance scores on the WIDA ACCESS for ELLs who participated in a Push-in or Pull-out instructional program, an independent samples t-test was conducted. If Levene’s test for equality of variances is significant, report the statistics for the row equal variances not assumed with the altered degrees of freedom rounded to the nearest whole number. The level of significance was .05 with \( F(1,118)=2.08, p=.152 \) a t-test assuming homogeneous variances was calculated. The results of this test indicated that there was not a significant difference in weight observed between the two groups, as shown in Figure 5 \( t(-1.67)=118, p=.099 \), Pull-out \( (M=243.8; SD=61.5) \), and Push-In \( (M=261.33; SD=53.7) \). Therefore, I accepted the null hypothesis claiming there is not a significant difference in test scores. In the Kindergarten Overall portion of the test, students’ scores were not dependent on the type of instruction they received.

**Grade 1**

This sample set consisted of 60 first grade students from District A ages six and seven. District B had a sample set of 48 first grade students’ ages six and seven. The students were chosen at random and their identities concealed, while using the independent t-test to compare the scores from both districts. The gender, race, or language spoken at home of participants were not identified. This information was not taken into consideration when randomly selecting the data. Results of listening, speaking, reading, writing, and overall scores are as shown in Table 6.

**Null Hypothesis 1.** There is no significant difference between the performances of scores on the listening portion of the WIDA ACCESS for ELLs who participate in a Push-in or Pull-out instructional program.
**Alternative Hypothesis 1.** There is a significant difference between the performances of scores on the listening portion of the WIDA ACCESS for ELLs who participate in a Push-in or Pull-out instructional program.

Table 6

*Independent Sample T-Test between students receiving Pull-out resources and Push-In Resources*

**GRADE 1-LISTENING portion of the WIDA**

N= Number of students

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Mean</th>
<th>Std. Deviation</th>
<th>Std. Error Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pull-out Resources</td>
<td>60</td>
<td>299</td>
<td>24.9</td>
<td>3.21</td>
</tr>
<tr>
<td>Push-in Resources</td>
<td>48</td>
<td>297</td>
<td>19.2</td>
<td>2.77</td>
</tr>
</tbody>
</table>

Independent Samples Test

<table>
<thead>
<tr>
<th>Change in weight post-test</th>
<th>Equal variances assumed</th>
<th>t-test for Equality of Means</th>
</tr>
</thead>
<tbody>
<tr>
<td>F</td>
<td>.801</td>
<td>Sig. (.373)</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>.583</td>
<td>Mean Difference (2.40)</td>
</tr>
<tr>
<td>Std. Error Difference (4.36)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

In order to examine differences in performance scores on the WIDA ACCESS for ELLs who participate in a Push-in or Pull-out instructional program, an independent sample *t*-test was conducted. If Levene’s test for equality of variances is significant, report the statistics for the row equal variances not assumed with the altered degrees of freedom rounded to the nearest whole number. The level of significance was .05 with *F*(1,106)=.801, *p*=.373 a *t*-test assuming homogeneous variances was calculated. The results of this test indicated that there was not a significant difference in weight observed between the two groups, as shown in Figure 1,
\( t(551)=106, \ p=.583 \), Pull-out\((M=299; \ SD=24.9)\) and Push-In \((M=279; \ SD=19.2)\). Therefore, I accept the null hypothesis claiming there is not a significant difference in test scores.

Table 7

*Independent Sample T-Test between Students Receiving Pull-out resources and Push-In Resources*

**GRADE 1-Speaking portion of the WIDA**

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Mean</th>
<th>Std. Deviation</th>
<th>Std. Error Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pull-out Resources</td>
<td>60</td>
<td>336</td>
<td>52.6</td>
<td>6.79</td>
</tr>
<tr>
<td>Push-in Resources</td>
<td>48</td>
<td>343</td>
<td>59.8</td>
<td>8.63</td>
</tr>
</tbody>
</table>

**Independent Samples Test**

<table>
<thead>
<tr>
<th>TABLE 2</th>
<th>F</th>
<th>Sig.</th>
<th>t</th>
<th>df</th>
<th>Sig. (2-tailed)</th>
<th>Mean Difference</th>
<th>Std. Error Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Change in weight</td>
<td>2.32</td>
<td>.631</td>
<td>-.713</td>
<td>106</td>
<td>.477</td>
<td>-7.725</td>
<td>10.83</td>
</tr>
<tr>
<td>post-test</td>
<td>Equal variances not assumed</td>
<td>-.703</td>
<td>94.41</td>
<td>.484</td>
<td>-7.725</td>
<td>10.98</td>
<td></td>
</tr>
</tbody>
</table>

Null Hypothesis 2. There is no significant difference between the performances of scores on the speaking portion of the WIDA ACCESS for ELLs who participate in a Push-in or Pull-out instructional program.

Alternative Hypothesis 2. There is a significant difference between the performances of scores on the speaking portion of the WIDA ACCESS for ELLs who participate in a Push-in or Pull-out instructional program.

In order to examine differences in performance scores on the WIDA ACCESS for ELLs who participate in a Push-in or Pull-out instructional program an independent samples \( t \)-test was conducted. If Levene’s test for equality of variances is significant, report the statistics for the row
equal variances not assumed with the altered degrees of freedom rounded to the nearest whole number. The level of significance was .05 with $F(1,106)=2.32, p=.631$ a $t$-test assuming homogeneous variances was calculated. The results of this test indicated that there was not a significant difference in weight observed between the two groups, as shown in Figure 2, $t(-.713)=106, p=.477$, Pull-out ($M=336; SD=52.6$) and Push-In ($M=343; SD=59.8$). Therefore, I accepted the null hypothesis claiming there is not a significant difference in test scores.

Table 8

*Independent Sample T-Test between Students Receiving Pull-out Resources and Push-In Resources*

<table>
<thead>
<tr>
<th>GRADE 1-READING portion of the WIDA</th>
<th>N= Number of students</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pull-out Resources</td>
<td>N= 60 Mean= 280.5 Std. Deviation= 33.3 Std. Error Mean= 4.30</td>
</tr>
<tr>
<td>Push-in Resources</td>
<td>N= 48 Mean= 280.9 Std. Deviation= 14.9 Std. Error Mean= 2.15</td>
</tr>
</tbody>
</table>

*Independent Samples Test*

<table>
<thead>
<tr>
<th>TABLE 3</th>
<th>F</th>
<th>Sig.</th>
<th>t</th>
<th>df</th>
<th>Sig.(2-tailed)</th>
<th>Mean Difference</th>
<th>Std. Error Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Change in weight post-test</td>
<td>8.03</td>
<td>.006</td>
<td>-.080</td>
<td>106</td>
<td>.937</td>
<td>-.412</td>
<td>5.19</td>
</tr>
<tr>
<td>Equal variances assumed</td>
<td>8.03</td>
<td>.006</td>
<td>-.080</td>
<td>106</td>
<td>.937</td>
<td>-.412</td>
<td>5.19</td>
</tr>
<tr>
<td>Equal variances not assumed</td>
<td>-.086</td>
<td>.855</td>
<td>.932</td>
<td>.412</td>
<td>4.81</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Null Hypothesis 3- There is no significant difference between the performances of scores on the reading portion of the WIDA ACCESS for ELLs who participate in a Push-in or Pull-out instructional program.
**Alternative Hypothesis 3.** There is a significant difference between the performances of scores on the reading portion of the WIDA ACCESS for ELLs who participated in a Push-in or Pull-out instructional program.

In order to examine differences in performance scores on the WIDA ACCESS for ELLs who participate in a Push-in or Pull-out instructional program an independent samples t-test was conducted. If Levene’s test for equality of variances is significant, report the statistics for the row equal variances not assumed with the altered degrees of freedom rounded to the nearest whole number. The level of significance was .05 with $F(1,85.5)=8.03, p=.006$ a t-test not assuming homogeneous variances was calculated. The results of this test indicated that there was a significant difference in weight observed between the two groups, Figure 3 shows, $t(85.5)=.086 p=.932$, Pull-out ($M=280.5$ $SD=33.3$) and Push-In ($M=280.9$; $SD=14.9$). Therefore, I reject the null hypothesis claiming there is not a significant difference in test scores and accept the alternate hypothesis.

**Null Hypothesis 4.** There is no significant difference between the performances of scores on the writing portion of the WIDA ACCESS for ELLs who participate in a Push-in or Pull-out instructional program.

**Alternative Hypothesis 4.** There is a significant difference between the performances of scores on the writing portion of the WIDA ACCESS for ELLs who participate in a Push-in or Pull-out instructional program.

In order to examine differences in performance scores on the WIDA ACCESS for ELLs who participate in a Push-in or Pull-out instructional program an independent samples t-test was conducted. If Levene’s test for equality of variances is significant, report the statistics for the row
equal variances not assumed with the altered degrees of freedom rounded to the nearest whole number.

Table 9

*Independent Sample T-Test between Students Receiving Pull-out Resources and Push-In Resources*

<table>
<thead>
<tr>
<th>GRADE 1-WRITING portion of the WIDA</th>
<th>N= Number of students</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pull-out Resources</td>
<td>N= 60</td>
</tr>
<tr>
<td>Push-in Resources</td>
<td>N= 48</td>
</tr>
</tbody>
</table>

**Independent Samples Test**

<table>
<thead>
<tr>
<th>TABLE 4</th>
<th>Levene’s Test for Equality of Variances</th>
<th>t-test for Equality of Means</th>
</tr>
</thead>
<tbody>
<tr>
<td>F</td>
<td>Sig.</td>
<td>t</td>
</tr>
<tr>
<td>Change in weight post-test</td>
<td>Equal variances assumed</td>
<td>.423</td>
</tr>
<tr>
<td>Equal variances not assumed</td>
<td>2.85</td>
<td>95.06</td>
</tr>
</tbody>
</table>

(Writing-Grade 1)

The level of significance was .05 with $F(1,106)=.423$ $p=.517$ a $t$-test assuming homogeneous variances was calculated. The results of this test indicated that there was not a significant difference in weight observed between the two groups, $t(106)=2.89$, $p=.005$, Pull-out ($M=273.4; SD=18.2$) and Push-In ($M=286.6; SD=20.5$), which is seen in Figure 4. Therefore, I accept the null hypothesis that there is not a significant difference in test scores.

**Null Hypothesis 5.** There is no significant difference between the performances of scores on the overall score of the WIDA ACCESS for ELLs who participate in a Push-in or Pull-out instructional program.
Table 10

Independent Sample T-Test between Students Receiving Pull-out Resources and Push-In Resources

GRADE 1-Overall portion of the WIDA

<table>
<thead>
<tr>
<th>N= Number of students</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pull-out Resources</td>
</tr>
<tr>
<td>Push-in Resources</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Independent Samples Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Levene’s Test for Equality of Variances</td>
</tr>
<tr>
<td>----------------------------------------</td>
</tr>
<tr>
<td>TABLE 5</td>
</tr>
<tr>
<td>Change in Equal variances assumed weight post-test</td>
</tr>
<tr>
<td>Equal variances not assumed</td>
</tr>
</tbody>
</table>

(Overall-Grade 1)

**Alternative Hypothesis 5.** There is a significant difference between the performances of scores on the overall score portion of the WIDA ACCESS for ELLs who participate in a Push-in or Pull-out instructional program.

In order to examine differences in performance scores on the WIDA ACCESS for ELLs who participate in Push-in or Pull-out instructional programs, an independent samples t-test was conducted. If Levene’s test for equality of variances is significant, report the statistics for the row equal variances not assumed with the altered degrees of freedom rounded to the nearest whole number. The level of significance was .05 with $F(1,106)=2.35, p=.129$ a t-test assuming homogeneous variances was calculated. The results of this test indicated that there was not a significant difference in weight observed between the two groups, $t(106)=.497, p=.620$, Pull-out
(M=290.9; SD=23.8) and Push-In (M=288.9; SD=16.9), as seen in Figure 5. Therefore, I accepted the null hypothesis claiming there is not a significant difference in test scores.

Grade 2

Grade 2’s sample set consisted of 60 second grade students from District A and thirty students from District B ages six through eight. The students selected were chosen at random and their identities concealed, while using the independent t-test to compare the scores from both districts. It is not clear what the gender, race, or language spoken at home is. This information was not taken into consideration when randomly selecting the data. Results of listening, speaking, reading, writing, and overall scores are as follows:

Table 11

<table>
<thead>
<tr>
<th>Independent Sample T-Test between Students Receiving Pull-out Resources and Push-In Resources</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>GRADE 2-Listening portion of the WIDA</strong></td>
</tr>
<tr>
<td>N= Number of students</td>
</tr>
<tr>
<td>Pull-out Resources</td>
</tr>
<tr>
<td>N= 60</td>
</tr>
<tr>
<td>Mean= 319.7</td>
</tr>
<tr>
<td>Std. Deviation= 49.1</td>
</tr>
<tr>
<td>Std. Error Mean= 6.34</td>
</tr>
<tr>
<td>Push-in Resources</td>
</tr>
<tr>
<td>N= 30</td>
</tr>
<tr>
<td>Mean= 314.6</td>
</tr>
<tr>
<td>Std. Deviation= 15.0</td>
</tr>
<tr>
<td>Std. Error Mean= 2.74</td>
</tr>
</tbody>
</table>

Independent Samples Test

<table>
<thead>
<tr>
<th>TABLE 1</th>
<th>F</th>
<th>Sig.</th>
<th>t</th>
<th>df</th>
<th>Sig.(2-tailed)</th>
<th>Mean Difference</th>
<th>Std. Error Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Change in Equal variances assumed weight post-test</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Equal variances not assumed</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Listening-Grade 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sig.(2-tailed)</td>
<td>.587</td>
<td></td>
<td>.546</td>
<td>88</td>
<td></td>
<td>5.02</td>
<td>9.29</td>
</tr>
<tr>
<td>Sig.(2-tailed)</td>
<td>.470</td>
<td></td>
<td>.727</td>
<td>77</td>
<td></td>
<td>5.02</td>
<td>6.90</td>
</tr>
</tbody>
</table>


Null Hypothesis 1. There is no significant difference between the performances of scores on the listening portion of the WIDA ACCESS for ELLs who participate in a Push-in or Pull-out instructional program.

Alternative Hypothesis 1. There is a significant difference between the performances of scores on the listening portion of the WIDA ACCESS for ELLs who participate in a Push-in or Pull-out instructional program.

In order to examine differences in performance scores on the WIDA ACCESS for ELLs who participate in a Push-in or Pull-out instructional program an independent samples t-test was conducted. If Levene’s test for equality of variances is significant, report the statistics for the row equal variances not assumed with the altered degrees of freedom rounded to the nearest whole number. The level of significance was .05, with $F(1,88)=3.64 \ p=.060$ a t-test assuming homogeneous variances was calculated. The results of this test indicated that there was not a significant difference, by a small margin, in weight observed between the two groups, as shown in Figure 1, $t(88)=.546, \ p=.587$, Pull-out ($M=41.9; SD=6.34$) and Push-In ($M=314.6; SD=2.74$). Therefore, I accept the null hypothesis claiming there is a significant difference in test scores.

Null Hypothesis 2. There is no significant difference between the performances of scores on the reading portion of the WIDA ACCESS for ELLs who participate in a Push-in or Pull-out instructional program.

Alternative Hypothesis 2. There is a significant difference between the performances of scores on the reading portion of the WIDA ACCESS for ELLs who participate in a Push-in or Pull-out instructional program.
Table 12

Independent Sample T-Test between Students Receiving Pull-out Resources and Push-In Resources

GRADE 2-Speaking portion of the WIDA
N= Number of students

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Mean</th>
<th>Std. Deviation</th>
<th>Std. Error Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pull-out Resources</td>
<td>60</td>
<td>351.8</td>
<td>63.7</td>
<td>8.22</td>
</tr>
<tr>
<td>Push-in Resources</td>
<td>30</td>
<td>366.9</td>
<td>52.6</td>
<td>9.61</td>
</tr>
</tbody>
</table>

Independent Samples Test

<table>
<thead>
<tr>
<th></th>
<th>Levene’s Test for Equality of Variances</th>
<th>t-test for Equality of Means</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>F</td>
<td>Sig.</td>
</tr>
<tr>
<td>Change in weight</td>
<td>1.88</td>
<td>.174</td>
</tr>
<tr>
<td>post-test</td>
<td>-1.2</td>
<td></td>
</tr>
</tbody>
</table>

(Speaking-Grade 2)

In order to examine differences in performance scores on the WIDA ACCESS for ELLs who participate in a Push-in or Pull-out instructional program an independent samples t-test was conducted. If Levene’s test for equality of variances is significant, report the statistics for the row equal variances not assumed with the altered degrees of freedom rounded to the nearest whole number. The level of significance was .05 with $F(1,88)=1.88$, $p=.174$ a t-test assuming homogeneous variances was calculated. The results of this test indicated that there was not a significant difference in weight observed between the two groups, $t(88)=-1.13$, $p=.263$, Pull-out ($M=351.8$; $SD=63.7$) and Push-In ($M=366.9$; $SD=52.6$), which can be seen in Figure 2. Therefore, I accepted the null hypothesis claiming there is not a significant difference in test scores.
Table 13

*Independent Sample T-Test between students receiving Pull-out resources and Push-In resources*  
**GRADE 2-Reading portion of the WIDA**  
N= Number of students

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Mean</th>
<th>Std. Deviation</th>
<th>Std. Error Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pull-out Resources</td>
<td>60</td>
<td>306.4</td>
<td>35.7</td>
<td>4.61</td>
</tr>
<tr>
<td>Push-in Resources</td>
<td>30</td>
<td>304.5</td>
<td>10.9</td>
<td>2.00</td>
</tr>
</tbody>
</table>

**Independent Samples Test**

<table>
<thead>
<tr>
<th>Equal variances assumed</th>
<th>F</th>
<th>Sig.</th>
<th>t</th>
<th>df</th>
<th>Sig.(2-tailed)</th>
<th>Mean Difference</th>
<th>Std. Error Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Change in weight</td>
<td>3.248</td>
<td>.075</td>
<td>.289</td>
<td>88</td>
<td>.773</td>
<td>1.93</td>
<td>6.69</td>
</tr>
<tr>
<td>post-test</td>
<td></td>
<td></td>
<td>.385</td>
<td>77.8</td>
<td>.702</td>
<td>1.93</td>
<td>5.02</td>
</tr>
</tbody>
</table>

(Reading-Grade 2)

**Null Hypothesis 3.** There is no significant difference between the performances of scores on the reading portion of the WIDA ACCESS for ELLs who participate in a Push-in or Pull-out instructional program.

**Alternative Hypothesis 3.** There is a significant difference between the performances of scores on the reading portion of the WIDA ACCESS for ELLs who participate in a Push-in or Pull-out instructional program.

In order to examine differences in performance scores on the WIDA ACCESS for ELLs who participate in a Push-in or Pull-out instructional program an independent samples t-test was conducted. If Levene’s test for equality of variances is significant, report the statistics for the row equal variances not assumed with the altered degrees of freedom rounded to the nearest whole number. The level of significance was .05 with $F(1,88)=3.25 \ p=.075$ a t-test assuming homogeneous variances was calculated. The results of this test indicated that although slight,
there was not a significant difference in weight observed between the two groups, \( t(88) = .289, p = .773 \), Pull-out (\( M=306.4; SD=35.7 \)) and Push-In (\( M=304.5; SD=10.9 \)), as seen in Figure 3. Therefore, I accepted the null hypothesis claiming there is not a significant difference in test scores.

Table 14

Independent Sample T-Test between Students Receiving Pull-out Resources and Push-In Resources

<table>
<thead>
<tr>
<th>GRADE 2-Writing portion of the WIDA</th>
<th>N= Number of students</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pull-out Resources</td>
<td>60</td>
</tr>
<tr>
<td>Mean</td>
<td>294.8</td>
</tr>
<tr>
<td>Std. Deviation</td>
<td>23.3</td>
</tr>
<tr>
<td>Std. Error Mean</td>
<td>3.01</td>
</tr>
<tr>
<td>Push-in Resources</td>
<td>30</td>
</tr>
<tr>
<td>Mean</td>
<td>284.6</td>
</tr>
<tr>
<td>Std. Deviation</td>
<td>18.4</td>
</tr>
<tr>
<td>Std. Error Mean</td>
<td>3.36</td>
</tr>
</tbody>
</table>

Independent Samples Test

<table>
<thead>
<tr>
<th>TABLE 4</th>
<th>F</th>
<th>Sig.</th>
<th>t</th>
<th>df</th>
<th>Sig.(2-tailed)</th>
<th>Mean Difference</th>
<th>Std. Error Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Change in Equal variances assumed weight post-test</td>
<td>1.83</td>
<td>.179</td>
<td>2.07</td>
<td>88</td>
<td>.041</td>
<td>10.11</td>
<td>4.88</td>
</tr>
<tr>
<td>Equal variances not assumed</td>
<td>2.24</td>
<td>.028</td>
<td>71.7</td>
<td></td>
<td></td>
<td>10.1</td>
<td>4.51</td>
</tr>
</tbody>
</table>

Null Hypothesis 4. There is no significant difference between the performance scores on the writing portion of the WIDA ACCESS for ELLs who participate in a Push-in or Pull-out instructional program.

Alternative Hypothesis 4. There is a significant difference between the performances of scores on the writing portion of the WIDA ACCESS for ELLs who participate in a Push-in or Pull-out instructional program.

In order to examine differences in performance scores on the WIDA ACCESS for ELLs who participate in a Push-in or Pull-out instructional program an independent samples \( t \)-test was
conducted. If Levene’s test for equality of variances is significant, report the statistics for the row equal variances not assumed with the altered degrees of freedom rounded to the nearest whole number. The level of significance was .05 with \( F(1,88)=1.83 \ p=.179 \) a 1-test assuming homogeneous variances was calculated. The results of this test indicated that there was not a significant difference in weight observed between the two groups, as shown in Figure 4, \( t(88)=2.07, \ p=.041 \), Pull-out(\( M=294.8; \ SD=23.3 \)) and Push-In \( (M=284.6; \ SD=18.4) \). Therefore, I accepted the null hypothesis claiming there is not a significant difference in test scores.

Table 15

*Independent Sample T-Test between Students Receiving Pull-out Resources and Push-In Resources*

**GRADE 2-Overall portion of the WIDA**

N= Number of students

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Mean</th>
<th>Std. Deviation</th>
<th>Std. Error Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pull-out Resources</td>
<td>60</td>
<td>311.6</td>
<td>33.7</td>
<td>4.36</td>
</tr>
<tr>
<td>Push-in Resources</td>
<td>30</td>
<td>308.7</td>
<td>15.9</td>
<td>2.90</td>
</tr>
</tbody>
</table>

---

Independent Samples Test

<table>
<thead>
<tr>
<th></th>
<th>Levene’s Test for Equality of Variances</th>
<th>t-test for Equality of Means</th>
</tr>
</thead>
<tbody>
<tr>
<td>TABLE 5</td>
<td>F</td>
<td>Sig.</td>
</tr>
<tr>
<td>Change in weight</td>
<td>4.36</td>
<td>.040</td>
</tr>
<tr>
<td>(Overall-Grade 2)</td>
<td>post-test Equal variances not assumed</td>
<td>.557</td>
</tr>
</tbody>
</table>

Null Hypothesis 5- There is no significant difference between the performances of scores on the overall score of the WIDA ACCESS for ELLs who participate in a Push-in or Pull-out instructional program.
Alternative Hypothesis 5- There is a difference between the performances of scores on the overall score portion of the WIDA ACCESS for ELLs who participate in a Push-in or Pull-out instructional program.

In order to examine differences in performance scores on the WIDA ACCESS for ELLs who participate in a Push-in or Pull-out instructional program an independent samples $t$-test was conducted. If Levene’s test for equality of variances is significant, report the statistics for the row equal variances not assumed with the altered degrees of freedom rounded to the nearest whole number. The level of significance was .05 with $F(1,88)=4.63$, $p=.040$ a $t$-test not assuming homogeneous variances was calculated. The results of this test indicated that there was a slightly significant difference in weight observed between the two groups, $t(87.8)=.557$, $p=.579$, Pull-out ($M=311.6; SD=33.7$) and Push-In ($M=308.7; SD=15.9$), as Figure 5 shows that Pull-out had a slightly higher value of test scores over Push-in resources. Therefore, I rejected the null hypothesis claiming there is not a significant difference in test scores.

Grade 3

This sample set consists of 60 first grade students from District A. District B had a sample set of 48 first grade students ages 7 and 8. The students were chosen at random and their identities concealed. The independent $t$-test was used to compare the scores from both districts. It was not clear what the gender, race, or language spoken at home is. This information was not taken into consideration when randomly selecting the data. Results of listening, speaking, reading, writing, and overall scores are shown in Table 1.
Table 16

Independent Sample T-Test between Students Receiving Pull-out Resources and Push-In Resources

GRADE 3-Listening portion of the WIDA
N= Number of students

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Mean</th>
<th>Std. Deviation</th>
<th>Std. Error Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pull-out Resources</td>
<td>45</td>
<td>347.9</td>
<td>35.3</td>
<td>5.3</td>
</tr>
<tr>
<td>Push-in Resources</td>
<td>30</td>
<td>318.6</td>
<td>22.8</td>
<td>4.2</td>
</tr>
</tbody>
</table>

Independent Samples Test

<table>
<thead>
<tr>
<th>TABLE 1</th>
<th>F</th>
<th>Sig.</th>
<th>t</th>
<th>df</th>
<th>Sig (2-tailed)</th>
<th>Mean Difference</th>
<th>Std. Error Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Change in weight post-test Equal variances not assumed (Listening-Grade 3)</td>
<td>1.64</td>
<td>.205</td>
<td>4.02</td>
<td>73</td>
<td>.000</td>
<td>29.3</td>
<td>7.29</td>
</tr>
<tr>
<td></td>
<td>4.37</td>
<td>.000</td>
<td>72.9</td>
<td>.000</td>
<td>5.02</td>
<td>6.71</td>
<td></td>
</tr>
</tbody>
</table>

Null Hypothesis 1. There is no significant difference between the performances of scores on the listening portion of the WIDA ACCESS for ELLs who participate in a Push-in or Pull-out instructional program.

Alternative Hypothesis 1. There is a significant difference between the performances of scores on the listening portion of the WIDA ACCESS for ELLs who participate in a Push-in or Pull-out instructional program.

In order to examine differences in performance scores on the WIDA ACCESS for ELLs who participate in a Push-in or Pull-out instructional program an independent samples t-test was conducted. If Levene’s test for equality of variances is significant, report the statistics for the row equal variances not assumed with the altered degrees of freedom rounded to the nearest whole number. The level of significance was .05 with $F(1, 73) = 1.64, p = .205$ a t-test assuming homogeneous variances were calculated. The results of this test indicated that there was a
A slightly significant difference in weight observed between the two groups, \( t(73)=4.02, p=0.000 \), Pull-out \( (M=349.9; SD=35.3) \) and Push-In \( (M=318.6; SD=22.8) \), as Figure 1 shows. Therefore, I could accept the null hypothesis claiming there is not a significant difference in test scores.

Table 17

*Independent Sample T-Test between Students Receiving Pull-out Resources and Push-In Resources*

<table>
<thead>
<tr>
<th>GRADE 3-Speaking portion of the WIDA</th>
<th>N</th>
<th>Mean</th>
<th>Std. Deviation</th>
<th>Std. Error Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pull-out Resources</td>
<td>45</td>
<td>341.8</td>
<td>50.7</td>
<td>7.6</td>
</tr>
<tr>
<td>Push-in Resources</td>
<td>30</td>
<td>350.1</td>
<td>49.6</td>
<td>9.1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>TABLE 2</th>
<th>F</th>
<th>Sig.</th>
<th>( t )</th>
<th>( df )</th>
<th>Sig.(2-tailed)</th>
<th>Mean Difference</th>
<th>Std. Error Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Change in weight post-test</td>
<td>.085</td>
<td>.772</td>
<td>-.707</td>
<td>73</td>
<td>.482</td>
<td>-8.38</td>
<td>11.85</td>
</tr>
<tr>
<td>Equal variances assumed</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Equal variances not assumed</td>
<td>-.710</td>
<td>.632</td>
<td></td>
<td></td>
<td>.480</td>
<td>-8.38</td>
<td>11.79</td>
</tr>
</tbody>
</table>

*(Speaking-Grade 3)*

**Null Hypothesis 2.** There is no significant difference between the performances of scores on the speaking portion of the WIDA ACCESS for ELLs who participate in a Push-in or Pull-out instructional program.

**Alternative Hypothesis 2.** There is a significant difference between the performances of scores on the speaking portion of the WIDA ACCESS for ELLs who participate in a Push-in or Pull-out instructional program.

In order to examine differences in performance scores on the WIDA ACCESS for ELLs who participate in a Push-in or Pull-out instructional program an independent samples \( t \)-test was
conducted. If Levene’s test for equality of variances is significant, report the statistics for the row equal variances not assumed with the altered degrees of freedom rounded to the nearest whole number. The level of significance was .05 with $F(1,73)=.085$, $p=.772$ a $t$-test assuming homogeneous variances was calculated. The results of this test indicated that there was not a significant difference in weight observed between the two groups, as shown in Figure 2, $t(73)=-.707$, $p=.482$, Pull-out ($M=341.8; SD=50.7$) and Push-In ($M=350.1; SD=49.6$). Therefore, I accept the null hypothesis claiming there is not a significant difference in test scores.

**Null Hypothesis 3.** There is no significant difference between the performances of scores on the reading portion of the WIDA ACCESS for ELLs who participate in a Push-in or Pull-out instructional program.

**Alternative Hypothesis 3.** There is a significant difference between the performances of scores on the reading portion of the WIDA ACCESS for ELLs who participate in a Push-in or Pull-out instructional program.

In order to examine differences in performance scores on the WIDA ACCESS for ELLs who participate in a Push-in or Pull-out instructional program, an independent samples $t$-test was conducted. If Levene’s test for equality of variances is significant, report the statistics for the row equal variances not assumed with the altered degrees of freedom rounded to the nearest whole number. The level of significance was .05 with $F(1,73)=17.6$, $p=.000$ a $t$-test not assuming homogeneous variances was calculated. The results of this test indicated that there was a significant difference in weight observed between the two groups, $t(35.7)=5.68$, $p=.000$ Pull-out ($M=344.9; SD=25.8$) and Push-In ($M=266.6; SD=62.3$), Figure 3 displayed the differences in mean score where students that received Pull-out services received higher scores in reading on the WIDA ACCESS. Therefore, I rejected the null hypothesis claiming there is not a significant
difference in test scores and accept the alternate hypothesis stating there is a difference in test score of students who receive Push-in or Pull-out services.

Table 18

*Independent Sample T-Test between Students Receiving Pull-out Resources and Push-In Resources*

**GRADE 3-Reading portion of the WIDA**
N= Number of students

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Mean</th>
<th>Std. Deviation</th>
<th>Std. Error Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pull-out Resources</td>
<td>45</td>
<td>334.9</td>
<td>25.8</td>
<td>3.9</td>
</tr>
<tr>
<td>Push-in Resources</td>
<td>30</td>
<td>266.6</td>
<td>62.3</td>
<td>11.4</td>
</tr>
</tbody>
</table>

Independent Samples Test

<table>
<thead>
<tr>
<th></th>
<th>F</th>
<th>Sig.</th>
<th>t</th>
<th>df</th>
<th>Sig.(2-tailed)</th>
<th>Mean Difference</th>
<th>Std. Error Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Change in weight</td>
<td>17.6</td>
<td>.000</td>
<td>6.57</td>
<td>73</td>
<td>.000</td>
<td>68.3</td>
<td>10.4</td>
</tr>
<tr>
<td>(Equal variances assumed)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Post-test</td>
<td>5.68</td>
<td>.000</td>
<td>35.7</td>
<td></td>
<td>.000</td>
<td>68.3</td>
<td>12.0</td>
</tr>
<tr>
<td>(Equal variances not assumed)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Null Hypothesis 4.** There is no significant difference between the performances of scores on the writing portion of the WIDA ACCESS for ELLs who participate in a Push-in or Pull-out instructional program.

**Alternative Hypothesis 4.** There is a significant difference between the performances of scores on the writing portion of the WIDA ACCESS for ELLs who participate in a Push-in or Pull-out instructional program.

In order to examine differences in performance scores on the WIDA ACCESS for ELLs who participate in a Push-in or Pull-out instructional program, an independent samples *t*-test was conducted. If Levene’s test for equality of variances is significant, report the statistics for the row.
equal variances not assumed with the altered degrees of freedom rounded to the nearest whole number. The level of significance was .05 with \( F(1,73) = 12.7, p = .001 \) a \( t \)-test not assuming homogeneous variances was calculated.

Table 19

*Independent Sample T-Test between Students Receiving Pull-out Resources and Push-In Resources*

<table>
<thead>
<tr>
<th>GRADE 3-Writing portion of the WIDA</th>
<th>N= Number of students</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pull-out Resources</td>
<td>45</td>
</tr>
<tr>
<td>Push-in Resources</td>
<td>30</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Mean</th>
<th>Std. Deviation</th>
<th>Std. Error Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pull-out Resources</td>
<td>45</td>
<td>346.7</td>
<td>27.3</td>
<td>4.1</td>
</tr>
<tr>
<td>Push-in Resources</td>
<td>30</td>
<td>281.2</td>
<td>61.2</td>
<td>11.2</td>
</tr>
</tbody>
</table>

Independent Samples Test

<table>
<thead>
<tr>
<th>TABLE 4</th>
<th>F</th>
<th>Sig.</th>
<th>( t )</th>
<th>df</th>
<th>Sig.(2-tailed)</th>
<th>Mean Difference</th>
<th>Std. Error Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Change in weight post-test</td>
<td>12.7</td>
<td>.001</td>
<td>6.30</td>
<td>73</td>
<td>.000</td>
<td>65.5</td>
<td>10.4</td>
</tr>
<tr>
<td>Equal variances not assumed</td>
<td>5.50</td>
<td>.000</td>
<td>36.8</td>
<td></td>
<td></td>
<td>65.5</td>
<td>11.9</td>
</tr>
</tbody>
</table>

The results of this test indicated that there was a significant difference in weight observed between the two groups, \( t(36.8) = 5.5, p = .000 \) Pull-out \( (M=346.7; SD=27.3) \), and Push-In \( (M=281.2; SD=61.2) \). Here, Figure 4 shows that Pull-out resources were more beneficial to the students. Therefore, it was possible to reject the null hypothesis claiming that there is not a significant difference in test scores and to accept the alternate hypothesis stating there is a difference in the test scores of students who received Push-in or Pull-out services.

**Null Hypothesis 5.** There is no significant difference between the performances of scores on the overall score of the WIDA ACCESS for ELLs who participate in a Push-in or Pull-out instructional program.
Table 20

*Independent Sample T-Test between Students Receiving Pull-out Resources and Push-In Resources*

**GRADE 3-Overall portion of the WIDA**

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Mean</th>
<th>Std. Deviation</th>
<th>Std. Error Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pull-out Resources</td>
<td>45</td>
<td>342.4</td>
<td>26.3</td>
<td>4.1</td>
</tr>
<tr>
<td>Push-in Resources</td>
<td>30</td>
<td>290.1</td>
<td>46.8</td>
<td>11.2</td>
</tr>
</tbody>
</table>

**Independent Samples Test**

<table>
<thead>
<tr>
<th></th>
<th>F</th>
<th>Sig.</th>
<th>t</th>
<th>df</th>
<th>Sig.(2-tailed)</th>
<th>Mean Difference</th>
<th>Std. Error Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Change in weight</td>
<td>6.91</td>
<td>.010</td>
<td>6.18</td>
<td>73</td>
<td>.000</td>
<td>52.3</td>
<td>8.45</td>
</tr>
<tr>
<td>(Overall-Grade 3)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**TABLE 5**

**Alternative Hypothesis 5.** There is a difference between the performances of scores on the overall score portion of the WIDA ACCESS for ELLs who participate in a Push-in or Pull-out instructional program.

In order to examine differences in performance scores on the WIDA ACCESS for ELLs who participated in a Push-in or Pull-out instructional program, an independent samples *t*-test was conducted. If Levene’s test for equality of variances is significant, report the statistics for the row equal variances not assumed with the altered degrees of freedom rounded to the nearest whole number. The level of significance was .05, with *F*(1,73)=6.91, *p*=.010 a *t*-test, not assuming homogeneous variances, was calculated. The results of this test indicated that there was a significant difference in weight observed between the two groups, *t*(41.3)=5.56, *p*=.000, Pull-out (*M*=342.4; *SD*=26.3) and Push-In (*M*=290.1; *SD*=46.8), as shown in Figure 5. Therefore, it was possible to reject the null hypothesis claiming there is not a significant
difference in test scores and to accept the alternate hypothesis stating there is a difference in the overall test scores of students who receive Push-in or Pull-out services.

Summary

Kindergarten outcomes showed that results were significant in the areas of listening and speaking, revealing that Push-in services most benefitted the students at that time. In turn, the results were not significant in reading, writing, and overall score. Grade 1 showed that listening, speaking, writing, and overall score were not significant. Grade 1 only showed significance in the area of reading where Push-in gave the students a better outcome on their mean scores. In the areas of listening, speaking, reading, and writing Grade 2 results were not significant although the overall score was significant where students received Pull-out services. Finally, Grade 3 was significant in reading, writing, and overall indicating that students benefitted the most when they received Pull-out services. Results were not significant in the listening and speaking portion of the test for Grade 3. The findings of the current study imply that perhaps neither program is independent of each other, and possible require a collaboration of the two types of instruction in order to best service the students.
CHAPTER 5
INTERPRETATION AND RECOMMENDATIONS

The results of this study have been reported in Chapter 4. Chapter 5 will include a summary of the study, an overview of the problem, a discussion of the purpose statement, the research questions, and the review of the methodology. This chapter will also include major findings, finally concluding with implications for further actions, a summary of the study’s key ideas, and recommendations for further research.

Overview of the Problem

As stated in Chapter 1, the number of students who receive ESL services is constantly rising in the United States. In addition to the population increase, state mandates in education have also increased school accountability. The increasing population of ESL students has sparked a discussion among educators as to which method is the best to teach these students. As the growth of the ELL population continues to outpace the growth of the PK–12 population and ELLs continue to score poorly across the content areas, it will be important for states to fully consider ELLs when implementing their reform plans (August, Estrada, & Boyle, 2012).

Although the review of the literature presented both the pros and cons of the Push-in and Pull-out methods, there is very little literature that provides a comparative study of the two. In order to best serve these students, we need to determine which model is the most beneficial for maximum gains in language acquisition.

Purpose of the Study/ Research Questions

The purpose of this study was to determine the most effective ESL intervention model, comparing it to the effect of the Push-in or Pull-out models of students in the four domains:
listening, speaking, reading, writing, and overall in grades kindergarten to three. The following question guided this clinical research study:

Do students who participate in ESL Push-in classes benefit more than students who receive Pull-out ESL classes based on their WIDA ACCESS scores?

The hypotheses were as follows:

**Null Hypothesis 1.** There is no significant difference between the performance scores on the listening portion of the WIDA ACCESS for ELLs who participate in a Push-in or Pull-out instructional program.

**Alternative Hypothesis 1.** There is a significant difference between the performances of scores on the listening portion of the WIDA ACCESS for ELLs who participate in a Push-in or Pull-out instructional program.

**Null Hypothesis 2.** There is no significant difference between the performances of scores on the reading portion of the WIDA ACCESS for ELLs who participate in a Push-in or Pull-out instructional program.

**Alternative Hypothesis 2.** There is a significant difference between the performances of scores on the reading portion of the WIDA ACCESS for ELLs who participate in a Push-in or Pull-out instructional program.

**Null Hypothesis 3.** There is no significant difference between the performances of scores on the speaking portion of the WIDA ACCESS for ELLs who participate in a Push-in or Pull-out instructional program.

**Alternative Hypothesis 3.** There is a significant difference between the performances of scores on the speaking portion of the WIDA ACCESS for ELLs who participate in a Push-in or Pull-out instructional program.
Null Hypothesis 4. There is no significant difference between the performances of scores on the writing portion of the WIDA ACCESS for ELLs who participate in a Push-in or Pull-out instructional program.

Alternative Hypothesis 4. There is a significant difference between the performances of scores on the writing portion of the WIDA ACCESS for ELLs who participate in a Push-in or Pull-out instructional program.

Null Hypothesis 5. There is no significant difference between the performance scores on the overall score of the WIDA ACCESS for ELLs who participate in a Push-in or Pull-out instructional program.

Alternative Hypothesis 5. There is a significant difference between the performances of scores on the overall score portion of the WIDA ACCESS for ELLs who participate in a Push-in or Pull-out instructional program.

Review of the Methodology

A quantitative research design was used to measure the results of the WIDA ACCESS test of students in grades kindergarten to three, using data from the years 2013-2015. Random sampling, “a sampling process in which each element chosen from a population has an equal chance of being selected” (Abbott, 2011), was chosen to collect data from District A and District B. Students in grades kindergarten through third and ages five through nine were chosen at random to compare the effects of the teaching methodology. Students in District A were labeled as receiving Pull-out instruction, and students in District B were labeled as receiving Push-in instruction. Students were compared by using an independent sample t-test in SPSS (Statistical Package for the Social Sciences).
Findings

Kindergarten

The results of the grade kindergarten of District A and B showed that in the topics of listening and speaking there was a significant difference. Here, the findings supported rejecting the null hypothesis, stating that there was no difference in the scores of students who received Pull-out instruction and Push-in instruction. Accepting the alternate hypothesis tells us that there was a difference in scores between students who received different types of instruction. When comparing the two districts, the data confirmed the null hypothesis, showing that there is no difference in test scores on the reading, writing, and overall tests. A possible cause for this outcome could be that many students at the kindergarten level do not leave kindergarten reading and writing. Moreover, the literacy rates among students in America are sobering. For example, “Sixty-six percent of all U.S. fourth graders scored ‘below proficient’ on the 2013 National Assessment of Education Progress (NAEP) reading test, meaning that they are not reading at grade level. Even more alarming is the fact that among students from low-income backgrounds, 80% score below grade level in reading” (National Center for Education Statistics (NCES), n.d.).

Grade 1

Upon completion of the independent samples $t$-test, the results supported the null hypothesis, for Grade 1, stating there is no difference in scores among students who receive Push-in and Pull-out services in the areas of listening, speaking, writing, and overall skills. On the contrary, there was a significant difference in the scores of first graders in the area of reading. We rejected the null hypothesis and accepted the alternate hypothesis, stating there is a difference in scores among the students who received these different types of teaching methods.
Grade 2

Evaluation of the outcomes of Grade 2 students showed that there was no significant difference in the areas of listening, reading, writing, and speaking. Students in these areas did not show a difference in scores relating to the type of services they received. Additionally, for the area of the overall score, where listening, speaking, reading, and writing are weighted, we found a small but significant difference showing that there is a difference in scores amongst the two districts. A possible reason for this outcome is that the overall score consists of a weighted score. To obtain the overall score WIDA takes the scores of the four domains and assigns numbers to each. The overall score uses 15% weight in listening and speaking, and 35% in reading and writing.

Grade 3

Grade 3 results told a slightly different story. In the areas of listening and speaking, we found that there was no difference in students’ test scores based upon the types of services they receive.

When reviewing the areas of reading, writing, and overall score, the independent samples t-test showed that there was a difference in scores amongst students that received Push-in and Pull-out instruction:

Many students have strong learning needs and preferences that do not match traditional classroom environments (e.g., formal seating and bright lights), or traditional methods of teaching (e.g., standardized texts, teacher lectures, and extensive, independent seatwork). Young children—and at-risk readers in particular—tend to be global, tactile, and kinesthetic learners. These children prefer and do well in classrooms that allow for movement, have some comfortable seating and varied lighting, and enable them to work
with relative ease in different groupings. Most important, research strongly indicates that when students’ environmental preferences are met, they are more likely to associate reading with pleasure, to read for longer periods, and, overall, to achieve higher scores. (Carbo, 2007)

The number of years that these students have been receiving the services and the amount that they are exposed to the English language could have put both of the districts on the same page. Perhaps this could be an effect of the sociocultural perspective on second language learning, based on the work of Vygotsky (1978), which highlights that all learning, including language learning, is based on social interaction.

**Conclusion**

The results of this research suggest that the outcomes of the students’ WIDA ACCESS show that students who receive Push-in and Pull-out services can both benefit students in grades kindergarten through third. The findings show that in some areas the type of instruction that the students received caused them to score differently on the tests. The independent samples t-test was conducted to find a difference in scores, but because it was a one way test and the researcher did not indicate which way, we do not know which is more beneficial, although the mean scores show that, in some areas, students had higher dependence on the services received. This study concurs with both opposing and supporting literature on each teaching method. It is important to reiterate that the study took place in two North Jersey districts where the schools were specifically dominated by a “non-diverse” student body. More variables should be considered in the analysis as stated in the recommendations for future research.
Recommendations for Future Research

The sample set used in this study was students chosen at random, but, because the information came from two school districts, many of the scores that have been provided could have been a result from the same students throughout the three-year span. Many students that receive services do not always exit out of the program within their first year. One outcome of this could be that there are other reasons besides the language barrier that the students are not progressing, such as a learning disability. A similar study using a larger sample set would make the outcomes of the $t$-test more useful. The larger sample set would provide the researcher with more random scores and, possibly, exclude duplicate participants in each grade level and domain comparing various districts throughout New Jersey. Also, a mixed method study Qualitative/Quantitative, using surveys indicating different types of curriculum and teaching styles of teachers adding observation of the students and teachers can provide more concrete results.

As the process of random sampling occurs, while remaining anonymous, weeding out scores based on if a student is eligible for Special Education services can result in different outcomes. Unfortunately with the WIDA ACCESS initial testing does not account for students who have a learning disability. A possibility is that a student can test into ESL because he does not speak, whereas the reason for this could be a type of speech impediment.

Further research on the student's history could be conducted to discover the language background of the student and the amount of English spoken outside of the school environment. The addition of qualitative research using surveys to learn the background of each student before analyzing their scores could also be useful. Comparing students with similar language settings could uncover better findings. In addition to finding the culture of the home life, surveys
including the educational career may sway the findings. In many districts, students are given the opportunity to attend English classes before they enter kindergarten. An essential key of the study could be based around the student’s history; possible venues to look into are Head Start Programs, language backgrounds, the students’ parents’ educational level, and their culture in general further exploring gender, age, and sex of the student.

In addition, surveying the students to see if they have a basis in their native language is important. Students who know the alphabet in their native language have a better chance of acquiring English: “At the beginning stages of acquiring L2, adults and adolescents with solid development in first language master basic interpersonal communication skills (BICS) faster than children after two to three years of exposure to second language. However, children achieve higher L2 proficiency in BICS with adults and adolescents typically retaining an accent” (Malarz & National Clearinghouse for Bilingual Education, 1991). Also, “When children's first-language development is discontinued before it is completed, they may experience negative cognitive effects in the development of L2” (Malarz & National Clearinghouse for Bilingual Education, 1991).

Observing the culture of the classroom and the curriculum that is being used could also provide valuable information. Students who are receiving Push-in instruction may only be learning the curriculum that is being taught in the mainstream classroom, while those who are being Pulled-out could be learning more of an acculturation type curriculum and missing what is being learned in the mainstream classroom. Teaching styles from the teachers of the students being observed can provide a different outcome; this also includes the classroom or group size of each.
Investigating the amount of education or professional development that the teachers receive in areas of like the English development of non-English speakers could be a necessity for further research. Educators exposed to this type of teaching style may better educate the students who are in the classroom and/or in small group instruction. Additionally, when the students are receiving Push-in, do the teachers team teach, or is it simply that the ESL teacher focuses on her students and following along with the classroom teachers curriculum.

Concluding Remarks

As educators, we all strive to provide the best education to our students. In addition to all the school rules and regulations that teachers have to keep, high-stakes testing standards, and local and state standards, we have students who are overburdened and cannot even speak the English language. In order to do this, we must have our students’ best intentions in mind, where we need to find their strengths and capitalize on them. “Acknowledging student strengths and developing rapport helps students recognize their own capabilities and realize how much you care” (Jewell, 2016). The results of this study did show that, in some instances, the type of instruction they receive does make an impact on the students learning. These suggestions for further investigation can possibly offer educators the knowledge to create the best type of programs to reach not only English Language Learning students but more students in general.
References


Education Week. (n.d.). Inclusion and the L.E.P. Student. Retrieved from


interview-opi-q1


McKeon, D. (1987). Different Types of ESL Programs. ERIC Digest, 4


Figure 1. Kindergarten listening.
Figure 2. Kindergarten speaking.
Figure 3. Kindergarten reading.
Figure 4. Kindergarten writing.
Figure 5. Kindergarten overall.
Figure 6. Grade 1 listening.
Figure 7. Grade 1 speaking.
Figure 8. Grade 1 reading.
Figure 9. Grade 1 writing.
Figure 10. Grade 1 overall.
Figure 11. Grade 2 listening.
Figure 12. Grade 2 speaking.
Figure 13. Grade 2 reading.
Figure 14. Grade 2 writing.
Figure 15. Grade 2 overall.
Figure 16. Grade 3 listening.
Figure 17. Grade 3 speaking.
Figure 18. Grade 3 reading.
Figure 19. Grade 3 writing.
Figure 20. Grade 3 overall.