A Quantitative Causal Comparison of State Standardized Test Scores in Cotaught and Self-Contained ELL Content Classes

William Aaron Zimkowski

Dissertation Submitted to the Doctoral Program

of American College of Education

in partial fulfillment of the requirements of the degree of

Doctor of Education in Second Language Instruction

June 2024

A Quantitative Causal Comparison of State Standardized Test Scores in Cotaught and Self-Contained ELL Content Classes

William Aaron Zimkowski

Approved by:

Dissertation Chair: Wes Anthony, PhD

Committee Member: David Mapp, EdD

Abstract

English language learners (ELLs) are a growing population in public schools across the United States. Achievement gaps between ELLs and English-speaking peers continue to widen despite receiving more resources and support. The problem is that ELL populations continue to fall behind each year due to poor English language development services, and they never develop the academic language needed for success in school. The purpose of this quantitative causal comparative study was to determine if there was a difference in standardized test scores depending on ELL placement in self-contained or cotaught content classes in Virginia. Current literature does not compare ELL classroom placement with academic performance on standardized tests. Vygotsky's social development theory served as the theoretical framework for the study because it applies to both language acquisition and content knowledge. Research questions and hypotheses sought to address if a significant difference existed in standardized test scores depending on ELL placement in either cotaught or self-contained classrooms. The null hypothesis stated there was no statistically significant difference in standardized test scores depending on ELL classroom placement. This study rejected the null hypothesis and found a statistically significant difference in ELL assessment scores in self-contained classrooms. In a large public school district in Northern Virginia, a random sample of 376 (n-376) students from archival data focused on test scores from the 2022–2023 school year.

Keywords: ELLs, cotaught, self-contained, sociocultural development, language, SOL, achievement, content, acquisition, academics, standards, environment, classroom, learner, students, English, history, math, ESOL

Dedication

This research project is dedicated to my wonderful and amazing family, wife, and children. My wonderful wife, Jennifer, has supported me throughout this process and constantly encouraged me during long weekends and countless frustrations. Without her support, this research project could not have taken place. Her love and guidance through internal and external challenges have made this process a manageable and transformative point in my life that I will carry throughout every endeavor. We have always grown together and are now looking forward to the next chapter in our lives. My mother, Joan, has supported me throughout my academic journey, career, and life over the last 38 years. Without her, none of this would have been possible.

This research is dedicated to my children, Claire, Violet, and Ethan, who have given me the determination to finish this project while overcoming the tendency to procrastinate and give up when facing financial difficulty. They have given me the strength needed to overcome many challenges and develop consistency within my work and life. I want to model the determination, grit, and reliance they will need to achieve their goals in life and know that a misguided high school student can achieve high academic standing and success. Through persistence and determination, anything is possible.

Acknowledgments

I want to acknowledge all the academic support I have had throughout the years as I have developed my language acquisition instructional practices. First, my experience at the University of Delaware as a learner and educator pushed me to become the language educator I am today. My language teachers opened the world and allowed me to explore the globe. These teachers recognized my abilities and guided me to language education, forever changing the course of my life. The English Language Institute opened the door for teaching language and highlighted the importance of comprehensive, intensive English programs I use today.

American College of Education educators have supported me throughout this project. The coursework and all the educators propelled me to become a confident instructional leader. My dissertation committee, Dr. Wes and Dr. Mapp, provided timely and complete feedback, making this dissertation possible. They have mentored me through this process, even during critical times when we barely made the deadline. Without their support, this research project would not be possible.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

Immigration worldwide is causing the need for public schools to provide better language acquisition services to more diverse learning groups. In the United States, multilingual learners are a growing population, and federal education legislation requires public schools to provide language acquisition services and support for English language learners (ELLs; Owens & Wells, 2021; Voss, 2022). This study examined the difference in the language learner's ability to become proficient in academic content measured by standardized test scores. There is a need to address achievement gaps and classroom placement in ELL populations and how student interactions impact achievement on standardized assessments.

This research study determined a statistically significant difference in ELL standardized mean test scores in mathematics and history in cotaught and self-contained classrooms. A statistically significant difference indicates a need for further investigation into the cause of the difference. Further research can also focus on which aspects of classroom placement cause ELL students to perform differently. The introduction to the study summarizes the problem background, problem statement, study purpose, significance of the study, research questions, hypotheses, theoretical framework, definitions of terms, assumptions, scope and delimitations, limitations, and a chapter summary.

Background of Problem

ELL populations are the subject of numerous qualitative and quantitative research efforts attempting to better serve diverse learning populations in public schools. Educators report a lack of effectiveness in preparing ELLs to meet the academic requirements in English (Szymanski & Lynch, 2020). ELLs must pass the exact graduation requirements in many states as English-speaking peers in English (Bond, 2020). Preparing ELLs to meet graduation

requirements is an issue that has grown over the past 2 decades, and practical solutions are needed to serve ELLs (Voss, 2021). ELL reading proficiency levels are 25%–30% lower than peers, and dropout rates are double for this group (Liping, 2021). By the eighth grade, 72% of ELLs score below basic proficiency levels in mathematics assessments (Swanson et al., 2021). Current literature examines different instructional approaches and social-emotional elements depending on the classroom model. Educators report a need for an understanding of the best practices to support language development and academic content (Rizzuto, 2017).

Current research identifies achievement gaps in ELL populations and teachers' preparedness to close these gaps (Owens & Wells, 2021). Researchers examine classroom settings and instructional models at great lengths to better service ELL populations (Rizzuto, 2017). Language acquisition also plays a vital role in the development of students, and not all services produce the same results. Public schools aim to have ELLs graduate on time regardless of language proficiency levels. In addition, ELLs are ultimately required to pass the exact graduation requirements in English and content classes in Virginia (Bond, 2020). Therefore, there is a need to determine which content service model works best for ELL populations developing proficiency in academic content.

Statement of Problem

The problem is that ELL populations continue to fall behind each year due to poor English language development services, and they never develop the academic language needed for success in school (Voss, 2021). This problem evolved from reports that show ELLs need to be more proficient in academic content classes (Consolidated State Performance Report, DOE, 2017). In addition, this problem has existed since the early 2000s, and educators do not feel professional development effectively supports ELLs in content classes (Kasper & Kellerman,

1997; Rizzuto, 2017). While ELLs also receive more resources, accommodations, and support to meet state requirements, academic achievement gaps in ELL populations widen each year, leading to 25% of ELLs dropping out of high school (Consolidated State Performance Report, DOE, 2017; Goodrich et al., 2021). ELLs in Virginia high schools experience similar challenges. Therefore, this problem remains an essential topic of study in educational research.

Castrillon (2017) showed that how ELLs communicate with educators and peers impacts their ability to access learning material. ELLs actively engaged in these communications score higher on standardized tests (Snow, 2016). Public schools use various combinations of service models to provide ELLs with academic support and language development (Sugarman, 2018). Many districts use English-dominate instructional models for content classes for speakers of other languages. Cotaught and self-contained content classes are the most common service models used in public high schools, and debate continues which model works best for content knowledge development (Case, 2015).

Research focused on how professional development and extensive training programs can close these achievement gaps and the potential causes of these achievement gaps (Li & Peters, 2020). The way ELLs interact with academic language varies significantly between these two service models, and current trends in the literature highlight the effectiveness of each service model and which is more beneficial for language development (Cole, 2018). This study contributes to filling this gap and has provided evidence that self-contained classes successfully prepare ELLs to meet graduation requirements. There is also a focus on different instructional models and their ability to close ELL achievement gaps (Mayer et al., 2018). Therefore, this research sought to identify if a difference existed in ELL standardized mean test courses in cotaught and self-contained classrooms.

Purpose of Study

The purpose of this quantitative causal comparative study was to determine if there was a difference in standardized test scores depending on ELL placement in self-contained or cotaught content classes in Virginia. A quantitative causal comparison study was the most suitable research approach because it highlighted the effects of the independent variables on the dependent variable (Taherdoost, 2022). This causal comparative study determined a significant difference in standardized mean test scores based on ELL classroom placement, either cotaught or self-contained, and identified statistically significant differences in the two groups (Fraekel & Wallen, 2012). Determining the relationship between the independent variables, cotaught and self-contained classrooms, and the dependent variable, standardized mean test scores, in ELL high school students was the goal of this study (Schenker & Rumrill, 2004).

This quantitative causal comparative study compared cotaught and self-contained classrooms, the independent variable, and standardized assessment scores, the dependent variable. The Virginia Standards of Learning assessments, the source of ELL student scores, have a reliability of >.80 overall and are a valid and reliable quantitative measure of learners' content understanding (Konold et al., 2018). An independent two-tailed *t*-test was used to determine whether the null hypothesis should be rejected or accepted (DeMoulin & Kritsonis, 2009). The independent and dependent variables supported this study by providing valid and reliable numerical data to explore the research questions and the research problem. Results of this study determined that classroom placement affects ELL standardized mean test scores and suggested that ELL placement is essential to academic success in public high schools.

The sample population of ELLs (n-356) was from a large public school district in Northern Virginia. By selecting a large random sample, this study sought to be generalizable to

other ELL populations (George, 2021). Data originated from archived district assessment scores of ELL students who took the Virginia Standards of Learning assessments during the 2022–2023 academic school year. To be included, the ELLs participating had to receive ELL services in their math or history classes, ensuring all research participants were getting language support in the content class. These data best supported this quantitative causal comparative study because they validated ELL knowledge of the content taught in the cotaught or self-contained classrooms.

Significance of Study

ELL student classroom placements have implications for all ELLs and ELL educators in public high schools in the United States. A difference in academic performance in high school ELL populations depending on class placement benefits educational communities for several reasons. Results of this study may give ELL educators an understanding of which classroom service model produces better results on standardized assessments and, perhaps, the role interactions play in ELLs developing content knowledge. ELLs who converse and discuss academic content with their peers perform better on standardized tests (De Jager, 2019). Communication in each classroom setting is different due to student grouping. A significant difference in test scores can be used to improve ELL instruction because educators can use the data to group students in the best learning environment for academic content.

This study highlights a potential connection between peer interactions and improved ELL content knowledge acquisition. This research discovered a significant difference in self-contained classrooms, and further inquiries can investigate the causes and effects. Results can also be shared with ELL administrators to implement schedule regulations and policies that best serve diverse ELL populations. The rationale for this study came from the need to improve ELL services to improve ELL content acquisition and prevent ELLs from not completing high school.

Research Questions

Two research questions guided this study to address the research problem and better serve ELL populations. The independent variables were cotaught and self-contained classrooms, and the dependent variable was standardized mean test scores. The mean scores resulted from the Virginia Standards of Learning Algebra I and World History standardized assessments administered in the 2022–2023 school year. Statistical analysis of comparing groups of test scores consisted of a two-tailed independent *t*-test to address the following research questions:

Research Question 1: What differences exist in standardized mathematics assessment scores between cotaught and self-contained ELL groups in a Northern Virginia school district?

Research Question 2: What differences exist in standardized history test scores between cotaught and self-contained ELL groups in a Northern Virginia school district?

Research Hypotheses

Two hypotheses were generated for each research question—a null hypothesis and an alternative hypothesis—in this study. Null hypothesis (Ho) assumes there is no significant difference in the mean standardized test scores (dependent variable) in the two groups, cotaught and self-contained classrooms (independent variables). In contrast, the alternative (Ha) hypothesis indicates a difference in mean standardized test scores in the two groups.

Ho1: There is no statistically significant difference in the standardized mathematics assessment mean scores between the cotaught and self-contained groups.

Ha1: There is a statistically significant difference in the standardized mathematics assessment mean scores between the cotaught and self-contained groups.

Ho2: There is no statistically significant difference in the standardized history assessment mean scores between the cotaught and self-contained groups.

Ha2: There is a statistically significant difference in the standardized history assessment mean scores between the cotaught and self-contained groups.

Theoretical Framework

Vygotsky's (1978) social development theory (SDT) states that a person's cognitive development is dependent on social interactions and the social environment where learning takes place (Ungvarsky, 2020). This framework was relevant to this study because social interactions for ELLs differ greatly depending on classroom placement. Vygotsky argued that there are stages of cultural development that learners go through to develop an understanding of learning material. All higher-order thinking functions also follow this pattern and begin with social interaction (Vygotsky, 1978). Thus, social interactions determine ELLs ability to access learning content and language.

Within social development theory, Vygotsky (1978) claimed that learning happens in the zone of proximal development (ZPD). This zone is the middle area between what learners can do independently and what they need assistance doing from peers or adults. All learning accrues in this space because the child can understand the input and manipulate output based on social interaction. Communication, both input and output, is needed for cognitive and second language development (Castrillón, 2017). In the same way, cognitive development is dependent on social interaction within the ZPD as is language acquisition. Social development theory (SDT) is an applicable theoretical framework because each high school ELL service model, cotaught or self-contained, attempts to arrange meaningful social interaction to develop students' language and content knowledge.

Social development theory framed this study because the research sought to identify significant differences in an academic content assessment that requires social interactions to develop proficiency. Looking for a difference in this way illustrates the importance of the interconnectedness of content acquisition and language development. Therefore, Vygotsky's SDT and ZPD theoretical framework worked effectively for this study because it equally applies to the language and academic content developed in each classroom service model. These theories also helped address the research questions and guided data analysis because ELLs' placement in different learning environments depends on their ability to communicate in English. Each classroom attempts to accommodate the ELLs' zone of proximal development to support the learning of academic content. The literature review following Chapter 1 further details social development theory and how it applies to this study.

Definition of Terms

This quantitative causal comparative study includes terms educators use that support ELL language development and academic achievement for continuity and understanding. These terms are used throughout this study, and the following definitions apply to this research:

Achievement gaps (AG) are discrepancies between various groups of students in the United States public school system (Howard, 2013).

Cotaught content classes are classes where content teachers and ELL teachers work together to adjust the curriculum to meet the needs of English language learners and language acquisition standards incorporated into content instruction (Sugarman, 2018).

English language learners (ELLs) are defined as students actively learning English while enrolled in public school and require language support to improve their academic performance, making high school more challenging (Bardack, 2010).

English to speakers of other languages (ESOL) is defined as classes and educators tasked with improving students' academic English skills that promote success in academic content classes (Sugarman, 2018).

Limited English proficiency (LEP) learners are limited in their ability to use English, making it difficult for them to participate in general education classrooms with native speakers (Olson, 2016).

Self-contained content ELL classes are defined as classes where dually certified teachers work to make the content more accessible for ELL students and also incorporate language acquisition (Sugarman, 2018).

Standards of Learning assessment (SOL) is a standardized academic content assessment in Virginia that determines whether students have met grade-level content proficiency levels (Bond, 2020).

Assumptions

Assumptions are beliefs accepted as accurate based on reason and logic but do not have proof or validation (Suresh, 2015). Students are placed according to their English language proficiency, and the class is appropriate for the needs of the student. Test scores from one academic year determined a significant difference in ELL test scores in self-contained classroom models. Assumptions were generalized based on these data points. In addition, the *t*-test statistical analysis has three basic assumptions. The sample population was evenly distributed; the variance among test scores within the population was equal; and the sample was randomly selected (Suter, 2011). These assumptions structured the statistical analysis of this research study.

Scope and Delimitations

The scope of a study is defined as the parameters the research study operates under and that frames the perspective of the research (Simon & Goes, 2013). The research focused on high school-aged ELLs receiving language and academic content support in Algebra I and World History I classes. Standardized test scores from the 2022–2023 school year determined that a difference existed in the cotaught and self-contained ELL classroom placements. The 2022–2023 school year was chosen to create a more accurate picture of student performance because it was the first regular school year since the COVID-19 pandemic school closures. The research for this study was sampled and conducted in a large Northern Virginia public school district.

Delimitations are the characteristics that develop from the boundaries within the scope of the study specific to the researcher and research project (Simon & Goes, 2013). First, this study consisted of standardized test scores from the 2022–2023 school. The standardized test scores of high school ELLs from Grades 9–11 enrolled in a supported math or history class were the focus of this study. Research was limited to 1 year of standardized test scores. This study was also limited to ELLs' standardized assessment scores regardless of their language proficiency level and grade-level placement. In addition, the scope was limited to districts with similar demographics and resources. These delimitations, combined with a large, randomized sample, increased the transferability of this study.

Limitations

Limitations in any study are conditions that are beyond the control of the researcher and are potential weaknesses to the generalizability of the study conclusions (Simon & Goes, 2013). There are two main limitations to this study. First, the district chosen for this study has many resources that are unavailable to other districts, and this could affect the generalizability of the

study. A large, randomized sample of 376 participants sought to alleviate this limitation and improve transferability. Secondly, English language proficiency level was not taken into consideration only classroom placement. This is due to the way archived data were gathered because the database that has the scores only marks if a student testing number is an ELL. Standardized tests scores do not take into consideration the English language proficiency level of the learner. Nevertheless, this research sought to determine ELLs' performance on standardized tests with a large, randomized sample.

These limitations affected the research study in several ways. Mainly, student's language acquisition levels affect their ability to understand the content on standardized tests. The research used standardized tests in math and history because they are less reliant on the student's English skills to alleviate this limitation. Personal bias always has the potential to influence research or the outcomes. Quantitative research seeks to minimize the influence of personal bias by using statistical analysis research methodology and randomized group sampling. Applying these quantitative strategies allows the results of this study to be more generalizable to the larger high school ELL population (Taherdoost, 2022).

Chapter Summary

ELLs are less likely to stay caught up in academic classes despite having more resources and support (Voss, 2021). Many public schools in the United States require ELLs to pass standardized assessments while developing their English language proficiency (Samson & Collins, 2012). Communication is at the core of language and content learning, and current research explores the many ways ELL students interact in the classroom (De Jager, 2019); however, there needs to be more focus on whether these methods of communication affect standardized test scores. The purpose of this quantitative causal comparative study was to

determine if there was a difference in standardized test scores depending on ELL placement in self-contained or cotaught content classes in Virginia.

Vygotsky's social development theory states that cognitive development relies on social interactions (Castrillón, 2017). This theoretical framework was used to guide the study because of the focus on ELL academic learning. ELL classroom placement maximizes the potential of learning academic content. The literature review outlined the theoretical framework of this study and current literature as they relate to the research problem, purpose, and questions.

Chapter 2: Literature Review

As English language learner (ELL) populations increase, they risk failing to meet graduation requirements and becoming long-term ELLs (Artigliere, 2019; Owens & Wells, 2021; Voss, 2021). There is a need for public schools to improve English language development services to meet these challenges and support language and academic development (Li et al., 2018). English language development service models differ significantly across public schools and districts (Goodrich et al., 2021). Specifically, high school students in Virginia, including ELLs, are expected to pass standardized tests in their core content classes to receive a high school diploma (Guy et al., 2000). The problem is that ELL populations continue to fall behind each year due to the poor English language development services, and they never develop the academic language needed for success in school.

This comparative causal quantitative study aimed to determine if a significant difference existed in ELL standardized test scores in cotaught or self-contained classrooms. The background of this problem evolved from the various service models public schools implement to deliver content and language acquisition services mandated by state and federal laws (Banse & Palacios, 2018). ELL classroom placement depends on many factors, and there is debate about whether inclusion or self-contained classroom models are more effective in developing language and academic content (Case, 2015). Whether the classroom setting affects the standardized test scores of ELLs and to what extent student-teacher and peer interactions support content mastery is unknown.

Current literature highlights each service model's benefits, drawbacks, and effectiveness and its various uses for ELLs (Cole, 2018; De Jager, 2019; Sugarman, 2018). There is significant research into the different types of communication in which ELLs engage and which is more

beneficial for language development (Cole, 2018). In addition, instructional models and their effectiveness in delivering language and content to ELLs combined with teacher and students' perceptions are the focal point of literature related to addressing this research problem (Howlett & Penner-Williams, 2020; Mayer et al., 2018; Zhang et al., 2021). Public school educators and teachers feel they need more preparation to meet ELLs' diverse needs, and current literature seeks to address and improve on this problem (Besterman et al., 2018).

The literature review features the theory supporting this study, reviews of current literature, and discusses of positives and negatives of different English language development and content classroom service models. This research compared the social interactions of ELLs based on their classroom placement and performance on standardized tests. Several central themes persistent in scholarly literature attempting to address the challenges facing ELLs were revealed (Lumbrears & Rupley, 2019). The literature review consists of the following sections: (a) teacher perceptions of ELLs, (b) teacher preparedness to meet ELL needs, (c) communicative language teaching, (d) student interaction, and (e) learning environment. A gap noted in the literature was comparison of English language development models.

Literature Search Strategy

A literature review was conducted through various digital library databases and search engines to identify recent, peer-reviewed, scholarly sources relevant to the research topic, problems, and questions. Relevant literature was obtained using the Educational Resource Information Center (ERIC), EBSC database, JSTORE, Google Scholar, SAGE Journals, and other academic databases provided by American College of Education. This research used peer-reviewed, scholarly articles within the last 5 years to investigate the research questions and the problem. In addition to searching scholarly databases, articles were found using references from

sources used in this research. References lists were used from similar studies and articles to locate research on language acquisition, communicative language teaching, ELL classroom models, and ELL achievement gaps. This research material came from books, government sources, professional organizations, and policy briefs.

The following search terms were used to search and explore the literature: *ELL, bilingual, bilingual instruction, self-contained, cotaught, achievement gaps, communicative language teaching, second language acquisition, K–12 program models, collaborative language learning, communicative language teaching, ESL, ESOL, program models, social development theory, zone of proximal development, ELL standardized testing,* and *literacy assessment.* These search terms are the most relevant to research problems and were used to locate relevant research in 2023.

Theoretical Framework

A theoretical framework must be established that aligns with the topic and current literature to address the research problem and research questions adequately. In Vygotsky's last work, *Thinking and Speech* (1934), he proposed sociocultural theory, which states that all intellectual learning and cultural values are based on one's prior knowledge and social interactions. Vygotsky's sociocultural theory originated in 1934 and explained intellectual development and its connection to social interactions. As a social psychologist, Vygotsky conducted most of his research in Soviet Russia. *Thinking and Speech* is a collection of essays written by Vygotsky to theorize how children develop cognitively and the optimal zone for where cognitive development occurs (Vygotsky, 1878). These theories were influenced by Jean Piaget and later further developed by him in the 1980s.

Vygotsky concluded that human development is a socially facilitated process and that social interactions are essential to academic development (Vygotsky, 1978). At the same time, Piaget (year) argued that social development is needed before any academic development in children. Sociocultural theory is regularly utilized in educational and language acquisition research because social interaction is required for linguistic and academic development (Vintan & Gallagher, 2019). This research examined if ELL classroom placement, with or without native-speaking peers, makes a difference in ELL standardized assessment scores. In other words, social interactions, which are essential for academic content development, are different depending on how ELLs are placed in content classes, and this difference could contribute to the problem of ELLs being less likely to meet graduation requirements (Voss, 2021).

The sociocultural theory related to this study because a critical difference between cotaught and self-contained classes is the holistic classroom experience and the social interactions that occur in all collaborative learning activities (Vintan & Gallagher, 2019). Abulhassan and Hamid (2021) concluded through quantitative measures that collaborative learning interactions are the preferred method of instruction for ELLs and ELL educators and are emphasized in all professional development training (Howlett & Penner-Williams, 2020). The sociocultural theory applies equally to language acquisition and content knowledge, assuming that content knowledge is obtained through interactions (Castrillón, 2017). Therefore, collaborative learning allows ELLs to develop language and content knowledge simultaneously.

The focus on social interaction for academic development is rooted in sociocultural theory. Collaborative learning interactions are needed for ELLs to develop oral proficiency in the target language and navigate difficult academic learning situations with peers (Abulhassan & Hamid, 2021). In addition, collaborative learning experiences allow students to discuss and

explore content through conversations with peers and teachers. Collaborative learning experiences are rooted in social development theory because social interactions are at the root of the learning process. Therefore, comparisons of content-area standardized test scores attained in different service models may provide evidence that ELLs' communication in the classroom affects their ability to pass standardized tests. The research questions specifically addressed this difference by comparing standardized test scores in cotaught and self-contained classes, where social interactions and differences in students' ability to develop content literacy are different.

Vygotsky (year) differentiated between learning and knowledge because knowledge can only be acquired through social interaction, and learning is the internalizing of concepts acquired after social interaction (Vintan & Gallagher, 2019). Organizing ELLs in different academic classroom service models with language acquisition support is an attempt by schools to support the development and delivery of academic content or knowledge (Castrillón, 2017). Classroom placement also is an attempt to bridge limited English proficiency and grade-level academic content to facilitate the learning process. The expectation is that students become proficient in academic content through social interaction (Vygotsky, 1978). By assigning cotaught or self-contained ELL classes, educators are attempting to accommodate lower English language proficiency levels by matching social interactions to the student’s specific needs.

The difference between learning, social interaction, and the internalization of academic concepts over time was explored by examining student performance depending on classroom placement. Zone of proximal development (ZDP) is where each person best absorbs and retains language acquisition and learning (Ungvarsky, 2020). The ZPD refers to the learners' ideal spot where learning academic content and language acquisition is most beneficial (De Jager, 2019). This area of optimal development is between what students can do on their own and where they

need linguistic and academic assistance (Castrillón, 2017; Vygotsky, 1978). ELL scores on academic standardized tests indicate which classroom setting best fits students' individual ZPD and enables optimal absorption of academic content.

Adapting grade-level content standards for students' language proficiency levels is an attempt by schools to match the classroom setting to ELLs' ZPD. Lessons, collaborative learning, and assessments are used for ELLs through various mediums and contexts to boost language acquisition and internalize academic content knowledge (Castrillón, 2019; Ungvarsky, 2020). In other words, ELLs' language proficiency, prior knowledge, and cultural identity all contribute to the learning process and influence the ZPD differently (Castrillón, 2019; Vintan & Gallagher, 2019). Educators seek to have ELLs understand and internalize academic content and language by placing ELLs in classes that meet their ZPD (Abulhassan & Hamid, 2021). By comparing ELLs' academic standardized test scores, educators determined that classroom settings and social interactions are successful in meeting students' ZPD and internalizing academic content.

Social interactions in various settings and contexts lead to the acquisition of language and the internalization of content knowledge. Vygotsky's sociocultural theory served as the overarching theoretical framework for this study (Castrillón, 2019). Social interactions in self-contained and cotaught classrooms attempt to accommodate ELLs' zone of proximal development. After all, academic learning cannot happen unless social interaction occurs, based on this theory (Vygotsky, 1934). In other words, social development theory states that cognitive understanding of educational content relies on social interaction and communication (Ungvarsky, 2020), which varies depending on ELLs' language proficiency levels. ELLs are improving their language ability and understanding of the content through dialogue and collaborative communication with classroom peers and teachers (Snow, 2014). The classroom setting is where

educators differentiate content to match English language proficiency levels because all higher-order thinking functions begin through social interactions (Vygotsky, 1978). Language use is essential in how people's thoughts and ideas are formed. This study examined how these social interactions affect ELLs' ability to meet graduation requirements (Li et al., 2018).

The two critical beliefs in social development theory are that learning is a progressive and linear process and that knowledge is developed through social interaction where learning occurs (Vygotsky, 1974). First, social interactions in various settings and contexts lead to the acquisition of language and cognitive development (Castrillón, 2019). The internalization of Vygotsky's social development theory served as the overarching theoretical framework for this study. Learning in both self-contained and cotaught classrooms attempts to accommodate ELLs' zone of proximal development. Vygotsky's social development theory is essential to education and language acquisition. This is because the expectation is that content and language can only be acquired through social interactions (Castrillón, 2017).

Vintan and Gallagher (2019) also used socio development theory to research the importance of collaboration between ELL educators and general education to improve ELL language and content acquisition. They argued that scaffolding used to accommodate language levels matches grade-level content to the ZPDs and concluded that districts should spend time and resources promoting collaborative learning for ELL populations (Vintan & Gallagher, 2019). Social development theory has several assumptions to support the hypothesis that social development proceeds and influences language acquisition and content development through school instruction (Vygotsky, 1934). Vygotsky assumed that through school instruction, concepts and words develop, academic concepts develop over time, and everyday concepts children hold cannot be generalized to the academic context. Any problem in learning processes

had to be studied (Vygotsky, 1934). These assumptions are considered while addressing problems regarding ELL achievement gaps (Voss, 2021). Each classroom setting works to develop academic content and scaffold instruction to improve understanding of academic content. ELL academic achievement can be studied using annual academic standardized test scores of ELLs.

This theory aligned with this study in several ways because of the difference in ELL communication depending on the classroom setting. For example, students in self-contained ELL classes are more likely to translate content and use their native language while engaged in collaborative learning. At the same time, integrated classes have more social interactions in the target language. This research examined these social interactions by comparing ELLs' scores on academic content assessments. The problem is that ELLs are not performing as well on standardized tests as their peers, and SDT states that social development comes before academic development (Castrillón, 2017; Voss, 2021). Therefore, ELLs' interactions with teachers and peers in the classroom are critical components of academic development.

The classroom environment and native speaker exposure differ greatly between classroom instructional models. This theory applied to this study because Vygotsky argued that these social interactions are the core of learning and linking prior knowledge and experience to develop new content knowledge and understanding (Vygotsky, 1978). By focusing on ELLs' performance on standardized yearly assessments, social development theory ZPD would imply that more social interactions with native speakers would benefit language acquisition; likewise, more collaborative learning experiences would develop content understanding. A significant difference in test scores would imply that Vygotsky's SDT and ZPD apply equally to language acquisition and mastering content standards. This sociocultural theory applies to second language

acquisition (Castrillón, 2017) and states cognitive development is dependent on social encounters which are needed for learning (Ungvarsky, 2020).

Other scholarly articles have used this theoretical framework when examining social interactions and cognitive development. De Jager (2019) used Vygotsky's (year) theory to explore students' perceptions of the teachers implementing an effective instruction model. Like this study, the SDT was appropriate because the differentiated instructional model requires teachers to adjust and present content to ELLs more comprehensibly. Because all cognitive learning develops through communication, the adjusted content caters to the individual student's ZPD. Students' responses indicated that when educators interwork, comprehensible input and collaborative learning are most beneficial for content understanding (De Jager, 2019). To further understand the research problem, a current literature review was conducted to build the foundation for understanding the research questions and addressing the research problems.

Research Literature Review

Over the years, public school systems have attempted to meet the needs of more diverse learning groups, including multilingual learners (Goodrich et al., 2021). The problem is that ELL populations continue to fall behind each year due to the poor English language development services and they never develop the academic language needed for success in school (Voss, 2021). In addition to this, educators are feeling less prepared to meet ELL needs and need to learn more about the language acquisition process and ways to support academic discourse in content areas (Artigliere, 2019; Miskill & Oliveira, 2019). The best way to address this problem is unknown because of the many variables that go into language acquisition and learning (Lertcharoenwanic, 2022). The COVID-19 pandemic made it clear that the learning environment is critical to academic development and language acquisition. ELL student achievement gaps

grew more during this time than in other groups (Guillén et al., 2020). Because of this, districts provide different service models that place ELLs in different settings to help close these gaps (Sugarman, 2018; Voss, 2021). These service models provide different learning environments and peer interaction, affecting how well ELLs can master content standards. To accommodate students' ZPD, educators focus on creating effective learning environments where ELLs work collaboratively to develop proficiency in academic content standards.

Achievement Gaps in Standardized Testing

Student ELL populations continue to grow, and countless variables determine their success in public high school (Voss, 2021). The current challenges that ELLs face in public high schools are not new. Over the last 24 years, content classes have been extremely challenging for ELLs, teachers, and administrators. Many experts feel systematic change is needed to properly service ELLs (Derwing et.al., 1999). Some progress has been made in improving performance in younger ELLs' performance in content classrooms; however, ELLs start school further behind and lack in academic skills needed to meet grade-level content standards (Goodrich et.al, 2020; Kieffer & Thompson, 2018). Goodrich et al. (2020) examined the differences in state and local ELL standardized assessment data and concluded that larger data sets lessen the achievement gap between ELLs and their peers; however, ELL service models and programs vary greatly throughout the country and these additional variables influence achievement gaps.

Standardized assessments, among other measures, are used to measure ELL achievement and determine graduation standards in Virginia and nationally (Kiefer & Thompson, 2018). There are also attempts nationally to reclassify ELLs by using different metrics; however, this is unlikely to exclude standardized test data (Bond, 2020). In addition, standardized test scores may not be a proper measure of language acquisition (Goodrich et al., 2020). Given the current

framework in which ELL success is measured on standardized tests, educators are forced to focus their content on the test items instead of overall student development (Lertcharoenwanich, 2020). Therefore, cotaught and self-contained service models in public schools that provide content and English language development are essential components of this problem.

Many researchers and school districts work to explore and lessen the problem of achievement gaps between monolingual and multilingual peers as this population grows (Besterman et al., 2018). Public schools need systematic change to support ELLs and a successful method of bridging content and language acquisition (Meskill & Oliveira, 2019). Properly preparing educators to meet this challenge requires examining each specific variable that contributes to the language and learning experience in the classroom.

Teacher Preparedness to Meet ELLs' Needs

Teachers play the most critical role in developing students' language and content knowledge to improve the academic success of all learners. A constant theme throughout ELL literature focuses on the teacher's training and preparedness to meet the diverse needs of their students (Szymanski & Lynch, 2020). In cotaught ELL service models, the teacher's planning, collaboration, and assistance with content and language standards improve the ELL classroom experience (Meskill & Oliveira, 2019). Teachers across the country are seeing an increase in ELLs in their content classes (Besterman et al., 2018); however, teachers feel less prepared to address the ELL challenges in the classroom (Szymanski & Lynch, 2020). Educators feel confident when professional development focuses on specific classroom strategies and when they better understand language acquisition standards (Song, 2016).

Districts have attempted to address these growing problems through professional development opportunities for teachers serving ELLs. Research showed extensive ELL

professional development and training programs improve ELL achievement (Li & Peters, 2020). Literature acknowledges the many factors contributing to language acquisition and the importance of the learning environment (Banse & Palacios, 2018; Bloor et al., 2023). Many professional developments for educators focus on ELL instruction and building a supportive learning environment. Research indicated that teacher enthusiasm and style are pivotal in ELL content and language acquisition (Hahl & Pietarila, 2021). Teachers find the standard professional development offered by districts time-consuming, uninformative, and overwhelming (Song, 2016). The investigation into the service model provides educators with confidence in knowing that the ELL classroom service models significantly differ in graduation requirements for ELLs.

Different Service Models Used for ELL Instruction

ELL language acquisition and content instruction vary in ways that satisfy federal guidelines, and ELLs receive English language development services minutes aligned with their English proficiency level (Artigliere, 2019). Federal guidelines provide this framework, and districts are free to develop different service models based on the needs of their student populations (Sugarman, 2018). There are standard service models that many districts use to deliver ELL language acquisition services and grade-level content instruction. The most common service models used in public schools in the United States are self-contained and integrated cotaught classroom arrangements or pullout models. Integrated cotaught classroom arrangements combine ELL with native speakers in content classes with language and content-certified teachers (Meskill & Oliveira, 2019). Self-contained classes consist of only ELLs, usually grouped by proficiency level, and instruction incorporates students' native language (Nahdiah, 2022). High school ELLs use a variety of service models in their master schedules depending on

their specific needs and learner populations. An overlap between service models may make it difficult to differentiate between them; however, ELL peer interactions and communication, which develop language acquisition, depend on the service model and the students in the class (Shen & Byfield, 2018; Sugarman, 2018).

Self-contained classes provide an ELL-only environment where ELL teachers build language and content instruction for ELLs separated from the general education classroom (Howlett & Penner-Williams, 2020). Self-contained classes group ELLs with other language learners based on their grade and English proficiency levels (Hackett et al., 2021). For example, self-contained classes receive various supports ranging from more bilingual support to a dually certified teacher (Britsch, 2021). Research indicated that ELL students who receive targeted interventions specific to their needs also improve their language development and academic success (Campbell & Filimon, 2018). In addition, using the native language can reduce the stress and anxiety that ELLs encounter when they do not understand the content (Nahdiah, 2022). Self-contained small group classes are the attempted intervention classroom setting allowing ELLs to develop language and content through a modified curriculum.

Cotaught classes group ELLs with other grade-level language learners and native English speakers in a general education classroom (Hackett et al., 2021). These classes usually have two teachers, one certified in content and one certified in English, to teach speakers of other languages (Hahl et al., 2021). To ensure success, cotaught classes require a close partnership between language and content teachers with shared planning time (Meskill & Oliveira, 2019). Through qualitative observations and interviews, Meskill and Oliveira (2019) determined that multimodal language and content-teaching partnerships can integrate content and language development through targeted lessons and formative feedback. ELLs also need a basic level of

comprehension and communicative ability to participate with their peers required for classroom participation (Shen & Byfield, 2018).

Communication Language Teaching in the Public-School Setting

ELL students communicate and interact with classmates differently in cotaught and self-contained service models. Oral communication and classroom discussion are essential for developing content literacy (Snow, 2014). In addition, content learning material is adapted to fit the learners' needs. Many teaching practices involve using learner support, like visual aids, to make content material more comprehensible for all students (De Jager, 2019). Comprehensible input and oral communication are essential components of communicative language teaching (CLT), which offers educators various ways to present language and content (Dos Santos, 2020).

Communicative language teaching is a practice based on several seminal language acquisition theories. Like the ZPD, interpersonal and cultural interactions are central to language development, a critical distinction from traditional language acquisition teaching models (De Santos, 2020). CLT aims to build language learners' ability to communicate in the target language and construct language not focused on grammar rules and translations (Adam & Berkessa, 2022). Adam and Berkessa also highlighted the benefits of CLT incorporating second language acquisition theory into a teaching approach for language and content where communication is the foundation of instruction.

Stephen Krashen's theory on second language acquisition serves as the foundation for CLT (Dan Guo, 2018). Krashen's theory revolves around five hypotheses to present a complete language acquisition model. These theories assert that language acquisition is subconscious, there is a difference between acquisition and the learned grammatical system, language acquisition happens predictably, and comprehensible input is the most crucial concept regarding

language acquisition (Krashen, 1982). Lastly, the affective filter hypothesis accounts for all the other factors, like motivation and self-esteem, that affect the language acquisition process. Therefore, the language acquisition process is heavily reliant on communication and input. (Krashen, 1982).

Krashens' (year) theories on language acquisition revolve around communication and how all these factors contribute to the language acquisition process. Similarly, sociocultural cultural theory, proposed by Vygotsky (1934), suggested social interaction is needed for academic content and development. Both theories play a pivotal role in accommodating students' English language proficiency, the ZPD, and content delivered in communicative ways (Dos Santos, 2020). Content classes for ELLs attempt to combine language acquisition services and grade- level content by providing differentiated instruction to ELLs.

The ZPD is the optimal area of learning between what learners can do independently and what they need in assistance from peers or the teacher (Howlett et al., 2020). School districts place ELLs in various settings to maximize language acquisition and academic development mutually (Sugarman, 2018). Therefore, ELLs' communication within their ZPD in each classroom may affect standardized test scores. In addition, CLT is a standard language acquisition teaching method in both service models (De Santos, 2020). Communicative language teaching methods can make grade-level content more accessible to ELLs.

Communicative language teaching is relevant to this study because both service models integrate these teaching practices in various ways. Language and content educators work to make the content more comprehensible for the language learner and address the many other affective variables that affect language acquisition (Çiftçi Ertürk & Özdemir, 2022; Hahl & Pietarila, 2021); however, teacher support and peer communication differ significantly between cotaught

and self-contained classrooms. Self-contained ELL content classes risk not supporting students' ZPD and widening achievement gaps (Owens & Wells, 2021). Cotaught classes risk not accommodating the many other factors that affect language acquisition, such as cognitive style, motivation, and learning style. (Banse & Palacios, 2018). The wide variety of variables that affect language acquisition and content development makes it difficult to determine which service models work better to boost ELL achievements on standardized assessments.

Current research highlights several drawbacks related to CLT in the academic environment. First, language proficiency levels vary greatly and determine learners' participation in CLT and various other factors, including self-efficiency (Dos Santos, 2020). Çiftci and Özcan (2021) concluded that traditional grammar-translation methods are better for teaching academic vocabulary and grammar. The pressure of standardized assessments and cultural differences all affect the implementation of CLT and the learners' receptiveness (Dos Santos, 2020). Nevertheless, Cole (2018), through a meta-analysis of research, concluded that language development outcomes were significantly larger in classes with peer-to-peer interactions. Therefore, peer-to-peer interactions, or collaborative learning opportunities, are essential in ELL academic development.

Collaborative ELL Learning Opportunities

Collaborative learning is essential to language and academic classrooms because language acquisition and content require communication to improve. Working with peers benefits language acquisition and content development in several ways. First, it allows ELLs to converse about content with their peers and engage in more complex thinking and problem-solving (Abulhassan & Hamid, 2021). Any classroom discourse where ELLs negotiate and discuss content promotes language development through interpersonal interactions and strengthens

fluency and oral intelligibility (Adem & Berkessa, 2022; Cole, 2018). In addition, Zhang et al. (2021) found that engaging and interactive collaborative learning opportunities benefit language development equally regardless of the target language. In other words, collaborative learning opportunities develop any language and academic understanding and are not specific to English language learners.

Collaborative learning opportunities allow all students to contribute to a group while interacting with authentic input and output and develop their comprehension of the material through peer interactions (Payant & Bell, 2022). Li and Peters (2020) found that teachers who taught ELL strategies using collaborative methods improved language development, content knowledge, and overall academic performance. These supports in collaborative learning environments are the educator's attempt to bridge the content to the learners' ZPD (Vintan & Gallagher, 2019).

Collaborative learning experiences are at the heart of social development theory because they allow students to explore problems through social interaction to complete a given task. Diverse groups of students work best collaboratively when they have a shared interest in finding a solution in each activity (Novitasari et al., 2018). Engaging ELLs in collaborative learning activities allows them to improve their oral fluency skills, have more interactions in the target language, and increase classroom participation (Abulhassan & Hamid, 2021). Therefore, collaborative learning experiences develop more communicative competence in ELL students, and ELL placement, and participation in cotaught or self-contained classrooms would alter collaborative learning experiences.

Collaborative Learning Opportunities in Cotaught Settings

In cotaught ELL content classes, ELLs are grouped by their grade level and placed in a general education class with a content-certified teacher and an English as a second language teacher (Sugarman, 2018). ELLs and general education students engage in various collaborative learning opportunities with their peers the same way they would in a general education class (Vintan & Gallagher, 2019). For example, students have group projects, presentations, and discussions throughout the school year. In integrated content classes, oral communication is the primary means to convey content and develop content literacy (Wright, 2016). These collaborative learning activities are communicative tasks that develop oral fluency, comprehension, and literacy skills (Payant & Bell, 2022).

There are potential drawbacks of collaborative learning opportunities in cotaught general education classrooms. Some ELLs and teachers need more benefits from collaborative learning opportunities (Abulhassan & Hamid, 2021). Students with less target language proficiency tend to shy away from working with peers, and other factors influence ELL participation in collaborative activities (Bense & Palacios, 2018). In addition, ELLs tend to resist learning if they do not feel respected or supported in their English language acquisition (Lumbrears & Rupley, 2019). A lack of planning time between the ESOL teacher and content teacher is one of the main barriers to collaborative learning activities (Vintan & Gallagher, 2019).

The peer-to-peer interactions in cotaught classes are between ELLs and native English speakers. Case (2015) examined these ELL and non-ELL interactions in detail through qualitative measures and drew several conclusions. The language barrier between ELLs and non-ELLs was a potential issue depending on the ELLs' proficiency and was prevalent in every activity. They found that the non-ELL viewed their role as aiding the ELL to understand the

content. Case (2015) described these interactions as two-fold. First, closing the gap is the educators' and the non-ELLs' attempt to make the content more comprehensible to the ELL through scaffolding and other supports. Then, interactions after the ELL understands the content improved participation in assignments and conversations that support content understanding.

Collaborative Learning Opportunities in Self-Contained Settings

Self-contained ELL content classes group ELLs together according to their English proficiency level, and the teacher is dually certified in the content and teaching English as a second language (Meskill & Oliveira, 2019). Spanish is the most common language spoken among ELLs in the United States (Pumpki et al., 2022). Therefore, collaborative learning opportunities use a combination of English and students' native language. Communication in the target language is needed for language development, and some research suggests it is academic literacy content that is needed (Krashen, 1982; Snow, 2014); however, studies have shown that native language use has a positive impact on affective factors that support language acquisition (Zhang et al., 2021). The mixture of the student's native language and target language creates more of a bilingual classroom setting where students discuss content in English and their native language. Nahdiah (2022) found that teachers and ELLs view the use of students' native language as a necessity for negotiating the meaning of academic content and explaining and clarifying complex concepts.

There are also potential drawbacks to collaborative learning opportunities in self-contained ELL content classes. Mainly, they do not expose the language learner to authentic, comprehensible input of the target language (Krashen, 1982). In addition, educators use more support and direct interventions during self-contained, collaborative learning because of the difficulty in assessing ELLs' progress in language development and content mastery (Nahdiah,

2022). Britsch and Shepardson (2021) concluded that ELL students relied heavily on teacher-created materials, could not synthesize the academic content, and merely copied information in a self-contained science class.

Peer-to-peer interactions in self-contained classroom settings are different in cotaught classes because all students are ELLs. Students' native language and cultural composition within the class would greatly determine and contribute to ELLs' willingness to engage in collaborative learning opportunities (Dos Santos, 2020). ELLs' native language, heritage, and culture amongst ELLs shape academic discourse within the self-contained classroom (Lumbrears & Rupley, 2019). However, it can be challenging to accommodate large classes of ELLs with many different cultures and languages.

Learning Environments Effects on ELL Performance

The learning environment affects ELLs' engagement and motivation and is essential to student success (Waxman et al., 2021). Waxman et al. (2021) found that students in effective learning environments feel more supported and have higher involvement. Educators draw from many different practices from education and language acquisition to build a positive and supportive learning environment for ELLs. Cultural, linguistic, and academic differences in ELL populations make learning the content more challenging, and educators seek to provide learning environments that are more engaging and supportive of each ELL's needs (Eichhorn et al., 2019). In addition, Mayer et al. (2018) found that educators significantly alter their expectations and support depending on the student's proficiency. Different educational strategies lead to consistency in both cotaught and self-contained classroom settings. Therefore, there are several ways research suggests developing effective learning environments in both cotaught and self-contained classes.

The Sheltered Instruction Observation Protocol, or SIOP, an instructional model, stands out in current practice, literature, and professional developments (Desjardins, 2020). The SIOP instructional model is a framework for delivering content-area standards to ELLs and improving academic language development (De Jager, 2019). SIOP aims to create an integrated, communicative, comprehensible approach to content instruction through teacher-led and student-to-student interactions. Instructional practices that cover lesson design and delivery comprise the framework of the SIOP model (Desjordins, 2020). Castrillón (2017) concluded that the SIOP model would be the best instructional model based on Vygotsky's sociocultural theory because educators attempt to create a learning environment conducive to the student's ZPD. The success of SIOP instructional strategies requires massive preparation and planning from ELL and content educators (De Jager, 2019). In addition to supporting ELLs when appropriately done, Yujuan (2021) determined that the SIOP model also improved new teachers' confidence in meeting the needs of language learners.

Banse and Palacios (2018) examined ELL student characteristics that shape the classroom learning environments and concluded that linguistic and culturally diverse classrooms should provide an inclusive classroom environment to promote ELL success. Similarly, Nahdiah (2022) concluded that these factors were essential for ELL academic success. Both service models can provide the proper learning environment in which ELLs thrive. SDT also states that effective measures that support learner growth can be incorporated into the classroom learning environment (Abulhassan & Hamid, 2021).

Cotaught and self-contained services models both take different approaches to achieve the same goal of developing ELL language and proficiency in grade-level content. Self-contained classes provide a rich and compelling learning environment that can support language

development (Zhang et al., 2021); however, these classes often need to be more concise in their content, making it easier for ELLs to pass standardized assessments (Desjordins, 2020). On the other hand, cotaught classes expose ELLs to the grade-level content and native speakers but may offer less support in English language development (Mayer et al., 2018). Many of these factors affect ELLs' learning environment and could affect ELL performance on standardized assessments, for example. Positive, comfortable, and effective learning environments foster communication, leading to higher student achievement (Yujuan Shi et al., 2021). Meanwhile, negative, unwelcoming, and ineffective learning environments hinder student achievement (De Jager, 2019). ELLs develop content literacy through discussion and social interactions between peers and teachers, which differ significantly between cotaught and self-contained ELL classes (Snow, 2014).

Chapter Summary

ELL research over 20 years highlights the need for more specific studies that can be easily generalized to the population as ELL populations continue to grow in the United States (Derwing et al., 1999). Given these themes throughout current literature, determining if a significant difference exists between the two main service models could improve ELL instruction and school experience. Both service models place ELLs in different social, cultural, and learning environments, providing collaborative learning opportunities, peer interactions, and teacher content presentation. Social development theory states that interactions and collaboration facilitate learning and communication, a critical aspect of constructing new knowledge (Ungvarsky, 2020).

Communicative language teaching also states that language development relies on interaction in the target language and comprehensible input (Dos Santos, 2020). Students’ social

and academic interactions differ depending on their classroom placement. ELLs are grouped academically and linguistically to encourage language and content development through classroom interactions. Grouped according to their communicative competence in the target language, ELLs receive academic content differently (Howlett & Penner-Williams, 2020). Current literature supports different ELL practices, but comparing the optimal content delivery to ELL populations is needed to improve academic performance.

Current literature highlights the potential benefits and pitfalls of collaborative learning opportunities in language learning. Research suggested that all ELLs should engage in collaborative learning opportunities that engage them in content discussions to develop target language ability and content literacy (Castrillón, 2019; Snow, 2014; Vintan & Gallagher, 2019). Research indicates teachers' ELL experiences and perceptions of collaborative learning vary drastically from highly positive and beneficial to harmful when English proficiency is low (Dos Santos, 2020; Payant & Bell, 2018); however, this research examined collaborative learning as it would naturally transpire in the classroom and determined a significant difference in these interactions. Results showing a significant difference could fill this literature gap and provide ELL educators with valuable information.

The learning environment is essential and affects many aspects of the educational and language acquisition process, including interpersonal communication between students and educators. Educators in both service models strive to create learning environments where ELLs feel they can engage in content and support language acquisition (Desjordins, 2020); however, the problem is that ELL populations continue to fall behind each year due to poor English language develop service and they never develop the academic language needed for success in school (Goodright et al., 2020). To investigate this problem, the research questions regarding

ELL classroom placement and the effect, if any, it has on standardized test scores, a theoretical framework using Vygotsky's (year) social development theory, explores classroom interactions among ELLs and their peers. The several ways schools provide language acquisition and content support services to ELLs influence student development. Content instruction and language acquisition culminate through classroom discussion. The more students communicate about the content, the better they perform on standardized assessments (Snow, 2014). Social development theory states that academic growth only improves after internalizing through the learning process (Vygotsky, 1978). By making the content accessible to ELLs' zone of proximal development, ELL educators are attempting to bridge this gap. A comparison of classroom models contributes to the literature by addressing how service models can affect ELLs' ability on standardized assessments.

Through comparing the effects of different classroom service models on content mastery, this study determined that the organization of language learners in different classroom models affects standardized test scores needed for graduation. Therefore, comparing ELL mastery of content standards through state assessment scores fills gaps in the literature and aligns with social development theory. A comparative quantitative research methodology analyzing grade-level content standardized test scores of ELLs in two different classroom models, cotaught and self-contained, fills this gap in the literature and provides valuable information about social interactions in ELL instruction in public schools. In the upcoming Chapter 3, the methodology and design used to explore the potential differences in ELLs' standardized test scores when they are served in cotaught and self-contained classrooms.

Chapter 3: Methodology

ELLs are a growing population that continues to face challenges in acquiring the academic language required to successfully complete graduation requirements and complete high school (Goodrich et al., 2021; Voss, 2021). The problem is that ELL populations continue to fall behind each year due to poor English language development services, and they never develop the academic language needed for success in school. The learning environment is an integrated aspect of language acquisition and students' academic success (Clark, 2021). There is a need to determine if schools are meeting the needs of diverse learning groups. ELL standardized assessments are used to measure ELL progress and achievement to determine if schools are meeting the needs of ELLs (Li et al., 2018). Public schools use a combination of cotaught and self-contained service models to deliver English language development and academic content to ELLs (Sugarman, 2018).

This quantitative causal comparative study investigated two research questions and hypotheses to address the research problem. The purpose of this quantitative causal comparative study was to determine if there was a difference in standardized test scores depending on ELL placement in self-contained or cotaught content classes in Virginia.

Research Question 1: What differences exist in standardized mathematics assessment mean scores between cotaught and self-contained ELL groups in a Northern Virginia school district?

Ho1: There is no statistically significant difference in the standardized mathematics assessment mean scores between the cotaught and self-contained groups.

Ha1: There is a statistically significant difference in the standardized mathematics assessment mean scores between the cotaught and self-contained groups.

Research Question 2: What differences exist in standardized history assessment mean scores between cotaught and self-contained ELL groups in a Northern Virginia school district?

Ho2: There is no statistically significant difference in the standardized history assessment mean scores between the cotaught and self-contained groups.

Ha2: There is a statistically significant difference in the standardized history assessment mean scores between the cotaught and self-contained groups.

By outlining the research process, a framework was established to increase the generalizability, reliability, and validity of this study. The following sections—research methodology, research design, the role of the researcher, research procedures, study population sample size, data instruments, data analysis, reliability and validity, and ethical procedures—-of this research are presented along with their rationale.

Research Methodology, Design, and Rationale

The purpose of this quantitative causal comparative study was to determine if there was a difference in standardized test scores depending on ELL placement in self-contained or cotaught content classes in Virginia. Using archived data from the 2022–2023 Virginia Standards of Learning standardized assessment data in Algebra I and World History I, the study identified a statistically significant difference in the mean test scores ELL high school students served in self-contained classrooms. This experimental methodology study sought to be generalizable and applicable to future ELL research. Results may contribute to educational leaders' conversations and decision-making about the most effective service models to meet the English and content acquisition needs of high school ELL students. This section establishes the rationale for the research methodology and design.

Methodology

This quantitative causal comparative study examined the differences in standardized mean test scores for groups of ELLs served in cotaught and self-contained content classes. The independent variables were the cotaught and self-contained classrooms, and the dependent variable was standardized test scores from the Virgina Standards of Learning assessments in mathematics and history. The rationale for the methodology selection came from a desire to improve ELL education overall and maximize ELL class time to improve standardized test pass rates and academic improvement. Passing these assessments is required for all high school students in Virginia. Therefore, a statistically significant difference in scores from different service models could help guide and determine ELL policy. Educators must maximize instructional time for language and content acquisition to boost ELL achievement and graduation rates.

Design

This study used a causal comparative research design to determine if there was a statistically significant difference in mean standardized test scores of ELLs served in self-contained and cotaught classrooms. This examined the independent variables of cotaught and self-contained classrooms and the dependent variable of standardized test scores. The rationale for the design aligns to the rationale for the methodology.

Addressing the research questions required statistical analysis of archived ELL standardized mean test scores. Independent *t*-test statistical analysis was conducted was conducted in Microsoft Excel to determine if the null hypothesis (Ho: $\mu 1 = \mu 2$) was accepted or rejected (DeMoulin & Kritsonis, 2009). The research used an alpha level of 0.005 to establish a significant difference in test scores depending on the service model (Spickhard, 2017). All

assumptions were upheld, and the statistical accuracy of the data was assured (DeMoulin & Kritsonis, 2009).

An independent two-tailed *t*-test was conducted. First, the numerical value of the dependent variable was calculated to establish the null and alternative hypotheses in numerical form (DeMoulin & Kritsonis, 2009). Next, an alpha level of .05 was used to establish a healthy confidence level in the results (Spickard, 2017). Then, the sampling distribution determined a z-normal distribution (DeMoulin & Kristonis, 2009). The decision rule was established to reject or accept the null hypothesis depending on whether the observational data fell at or beyond the critical numerical value.

Role of the Researcher

As a doctoral candidate, I took responsibility for and conducted this quantitative causal comparative study. My role was to determine if a significant difference in standardized test scores of two groups of ELLs served in cotaught and self-contained classroom service models (Tanenbaum et al., 2012). I adhered to strict confidentiality standards to ensure that research participants' identifying factors were not revealed. Using archived data and established statistical measures allowed me to manage outside variables that may have undermined the research process. I am an educator at the research site and have a vested interest in improving ELL instruction in the county; however, the research design allowed me to conduct research free of bias and adhere to the principles of *The Belmont Report* (1979).

I followed all principles of *The Belmont Report* (1979) to ensure research credibility and ethics. *The Belmont Report* (1979), born out of the need to make scientific research ethical, establishes that all research must meet specific criteria. First, I ensured that all participants were treated respectfully and autonomously (U.S. Department of Health & Human Services, 1979).

Second, I demonstrated that the research benefited research participants and contributed to the ELL research. Lastly, all research benefits and burdens were justly distributed among participants. My role was to strictly adhere to these standards before, during, and after the research process.

Research Procedures

Population and Sample Selection

The population for this research included ELLs in Grade 9–12 in a Northern Virginia public school district. Male and female participants were pursuing a standard high school diploma in Virginia's public schools. Only ELL students actively receiving English language development services and other academic supports were included in the data sets (George, 2021). ELL students were served either in cotaught or self-contained classrooms. Dually identified special education students were excluded from the data pool. Randomized sampling was used to ensure the research's validity, generalizability, and reliability (George, 2021). Site permission was obtained to access the archived data to conduct this research. The Data Analysis and Reporting Department approved this research request through the district office. An online application was used to submit the research design, purpose, and description. After review, permission was granted, and temporary login credentials were given to gather data and conduct research (Appendix A).

The research site had a total of 4,995 ELLs in several high schools during the 2022–2023 school year. This study used a sample size of 376 (n 376). A sample size was calculated with a 95% confidence level, confidence interval of 5%, and a standard deviation of 0.5. This sample size allowed more generalizability to the entire ELL population because of the high confidence level and low margin of error or standard deviation.

Archival Data

This study used archived data saved by a large school district in Northern Virginia of Standards of Learning end-of-course (SOL EOC) standardized assessment scores for high school ELLs. The SOL EOC data were gathered from the school district's data management system. The Data Analysis and Review (DART) system collects and stores all district assessment data and student records. Archived data require district written approval to access and receive login credentials. A written approval letter was obtained and attached in Appendix A. Login credentials were granted for the approved research timeframe.

After approval, data were gathered from the DART database. The database imported ELL SOL EOC assessment scores in Algebra I and World History I. Archived test scores were accessible through district computers or remotely. Parameters (population, sample size, assessment scores, etc.) were used to gather proper data. The DART system generated data reports containing the requested data points. For example, a random sample of test results was pulled from all 9–12 ELL high school students' SOL EOC assessment scores. After that, the SOL EOC assessment scores were organized with parameters, including the service delivery model used to implement instruction, the independent variable. Disaggregated data allowed an investigation of necessary subsections and differences of ELL populations with specific parameters (Tanenbaum et al., 2012). Test scores were randomized and compared for any significant difference based on the service delivery model used for instruction**.**

Virginia Standards of Learning End-of-Course Assessments

The data instruments used for this quantitative causal comparative study were the Standards of Learning end-of-course standardized assessments (SOL EOC). SOL EOC assessments are content class summative assessments that cover the entire curriculum taught in

English language arts, history, math, and science courses. Virginia Department of Education creates the SOL EOC assessments. According to Konold et al. (2018), the SOL EOC assessments are developed through a procedure that limits item bias and ensures content knowledge is being tested. SOL tests are administered annually to determine if students meet the state proficiency levels in content-level classes. This study focused on ELL SOL EOC standardized scores in math and history classes, the dependent variable in this study.

In alignment with the research question and problem, the SOL EOC assessment data provided the best measure to gauge ELLs' understanding of content. SOL EOC assessments are adaptive and alter difficulty based on students' previous responses. ELLs who struggle with the questions are given more straightforward questions to assess their content knowledge. A comparison of these scores allowed the research to demonstrate that a significant difference existed in the standardized mean test scores of ELLs who were served in a self-contained service model.

Instrument Validation

The SOL assessments are the state of Virginia's annual content-area assessments. ELLs are scored on a level from 0 to 600. Scores above 400 pass proficient and meet state graduation requirements. Scores above 500 are pass advanced. Scores from 375 to 399 are eligible for retakes. The SOL EOC assessments have a reliability of >.80 overall for each subgroup (Konold et al., 2018). According to the Virginia Department of Education (1999), SOL EOC assessments are valid and reliable in several ways. First, SOL assessments are criterion-referenced assessments developed by multiple parties and verified as valid and reliable through independent assessment specialists (Virginia Department of Education, 1999). In addition, each test item is vetted through a process to ensure the questions are accurate, target the standard being assessed,

and are free of bias. The Virginia Department of Education also used outside testing evaluation experts to improve the validity and reliability of the results so state-level educational decisions can be made using them (Ruff, 2019).

Data Collection

To address the research problem and questions, a request for approval to access the district's data management system or DART was approved. The DART system required district approval to access assessment data from the last 5 years. The district database, DART, houses all the district's data. Precisely for this research, DART holds SOL EOC assessment data, students' ELL status, and instructional methods used to service ELLs. First, DART was accessed, and reports for scores were generated using parameters established for the research. The DART system generated a score sheet that included all the SOL EOC ELL scores on the desired assessment. Then, a random sample was drawn from all the assessment scores and used in the independent two-tailed *t*-test.

Data were collected and downloaded into Excel spreadsheets after approval was given. The file was kept on an encrypted computer that is password protected and annually wiped clean. The data spreadsheet had no identifiable information to protect participants' identities. A spreadsheet containing SOL EOC grades for ELL students during the 2022–2023 academic school year. Only ELLs who participated in the SOL EOC assessment and met graduation requirement scores were used. SOL EOC test scores were organized by assessment. For example, there were two categories: algebra and world history. Data were sorted into two categories of cotaught and self-contained ELL classes to align with the independent variable. Each category had an equal number of participants.

Participant information was kept secure by removing all identifying information before statistical analysis. Depending on the service delivery model, the data sets or testing categories were labeled as cotaught or self-contained. Statistical analysis consisted of test scores and no other identifying information. Research data were secured on a district computer that is cleared each summer.

Data Analysis

To address the research questions and hypotheses, this research used a two-tailed independent *t*-test to determine if the null hypothesis should be accepted or rejected (Kataike et al., 2017). Data were collected from district archives and county assessment management software. After data collection, subjects were organized into two independent samples based on the independent variables, cotaught or self-contained classroom (DeMoulin & Kritsonis, 2009). The SOL EOC assessment cotaught class score ($\mu 1$) was compared to the SOL EOC assessment self-contained class score ($\mu 2$) to determine whether a significant difference existed in assessment data dependent on classroom service model placement (Kataike et al., 2017).

To address the problem in this study, Microsoft Excel was used to conduct a two-tailed independent *t*-test to determine if the null hypothesis (Ho: $\mu 1 = \mu 2$) should be accepted or if it should be rejected for the alternative (DeMoulin & Kritsonis, 2009). This research used a confidence level (p-level) of 0.005 to determine whether any significant difference existed between the two groups (Spickard, 2017). The Ho would be rejected if the observed value fell beyond the critical; otherwise, it would fail to be rejected. This study maintained that all *t*-test assumptions were kept and followed to avoid misinterpretation of the data (DeMoulin & Kritsonis, 2009).

An independent two-tailed *t*-test was conducted in the following manner. First, the numerical value of the dependent variable was calculated to establish the null and alternative hypotheses in numerical form (DeMoulin & Kritsonis, 2009). Next, an alpha level of .05 was used to establish a healthy confidence level in the results (Spickard, 2017). Then, the sampling distribution was a z-normal distribution (DeMoulin & Kristonis, 2009). The decision rule was then established to reject or accept the null hypothesis depending on whether the observational data fell at or beyond the critical level. This study rejected the null hypothesis and found a statistically significant difference in ELL assessment scores in self-contained classrooms.

Reliability and Validity

Validity and reliability are essential aspects of all research and focus on these aspects improved the robust nature of the current study, so results were generalizable to the high school ELL populations. Reliability and validity were strengthened in several ways for this quantitative causal comparative study. First, stratified sampling was used to organize data clusters into independent variable groups depending on classroom placement (George, 2021). This lowered the threats to reliability and validity through random sampling. Secondly, the dependent variable was state-standardized test scores that ensured validity because SOL scores are independently validated by state regulators and testing organizations. All assumptions for the two-tailed independent *t*-test were met, making the research valid and reliable while increasing the generalizability (DeMoulin & Kritsonis, 2009).

This quantitative causal comparative study had one potential threat to the validity, reliability, and generalizability. The district selected for the research is a large district with many more resources and support than other districts throughout the United States. These resources could affect students' performance on standardized test scores. In addition, the district's

resources for ELLs include free summer testing preparation programs, purchased computer programs, and educator support. Many of these variables threaten validity, reliability, and generalizability because other districts with fewer resources cannot support extra programs. In addition, these variables can directly influence ELL performance on the assessments; however, future research could pull the same data from different districts and compare the results. This further validated this research and improved its reliability.

Ethical Procedures

Ethics is an essential aspect of research and protects the privacy and safety of research participants while ensuring researchers conduct themselves honestly. *The Belmont Report* (1979) established that all research should be respectful, beneficial, and just for all participants. Participants were respected by keeping all their information confidential and removing any identifying information. All research data were kept confidential and secure to ensure all research subjects were treated with dignity and respect (Dhai, 2017). Determining if a difference existed between cotaught and self-contained classes was beneficial as it highlighted the potential importance of the ELL service models. A significant difference in how ELL classes are organized suggests that further research can pursue the reasons. In addition, a significant difference in SOL EOC scores indicate that one ELL class setting may be preferred. This study was just because it was generalizable to the larger ELL population in high schools due to the large sample size. No conflicts of interest or power differentials were related to this research.

Chapter Summary

Data collected determined a significant difference in SOL EOC ELL assessment scores and addressed the research questions. To conclude if there was a significant difference, a quantitative causal comparative design using two-tailed independent *t*-tests was selected for this

research. The population focused on high school-aged ELLs in Northern Virginia seeking a high school diploma. A random sample (n 376) boosted the validity of the results, allowing the conclusions to be generalized to the entire ELL population. Standardized test scores were used as the dependent variable for their reliability and validity, whereas the classroom service models were the independent variables. This study provided valuable insight into how an ELL classroom's service model influences content-area standardized test scores. There are very few ethical considerations due to using archived district data. The three principles of *The Belmont Report* were implemented throughout the research process. The results of this study are explored in the following chapter.

Chapter 4: Research Findings and Data Analysis Results

Public schools must find ways to maximize instructional time for English language learners (ELLs) to increase their ability to access content while acquiring academic English. Various ELL instructional models exist to provide language and academic support, including classroom placement (Sugarman, 2018). ELLs are a growing population, and academic achievement gaps between ELLs and English-speaking peers widen despite receiving academic and linguistic services in public schools (Olsen, 2014). The achievement gap widens as ELLs work through school, complicating their ability to meet graduation requirements and complete high school on time (Goodrich et al., 2021). To improve graduation rates and student success, there is a need to compare ELL service models to determine if ELL needs are met.

The problem is that ELL populations continue to fall behind each year due to poor English language development services, and they never develop the academic language needed for success in school (Olsen, 2014). The purpose of this quantitative causal comparative study was to determine if there was a difference in standardized test scores depending on ELL placement in self-contained or cotaught content classes in Virginia. Comparing 2022–2023 ELL Algebra I and World History I SOL scores and classroom placement was used to determine if a significant difference existed with the potential to improve ELL academic achievement.

SOL and classroom placement data were collected from a school district in Northern Virginia. The county prepared these data and delivered them through secured school email. The statistical analysis results guided the research discussion, addressing the research questions and hypotheses. Reliability and validity were addressed to solidify the study results' credibility, dependability, and generalizability.

Data Collection

This study used archived standardized assessment data from ELLs in Grades 9–12, which are required for graduation in Virginia (Ruff, 2019). Data were collected at the start of the 2023 school year. The test scores used as data for this research study represented Algebra I and World History I tests taken during the 2022–2023 school year. The data file comprised the entire district's Algebra I and World History I test scores. A total of 5,338 test scores were in the file, with a course code attached. Assessment data is processed and prepared by the district research department that manages all district data needs. A data file was prepared and sent after site approval. The anticipated data collection plan was consistent.

Data collected for this study came from a large school district in Northern Virginia with about 90,000 students, 27.7% of whom were ELL. At the end of the 2022–2023 school year, 71% of ELL students graduated on time. The district has more than 150 languages spoken by families and students (PWCS, n.d.). The overall ELL pass rate on the Algebra I SOL assessment was 60%, and the pass rate on the World History I SOL assessment was 28% (PWCS, n.d.). \

Table 1

Descriptive Statistics of the Sample Population

	Total Student Population	Graduation Rate	Algebra SOL Pass Rate	World History SOL Pass Rate
District	71.8%	95.6%	79%	66%
ELL	28.2%	71.3%	60%	28%

Data were collected from high school ELL students enrolled in Algebra I and World History I who took the end-of-course standardized assessment in the 2022–2023 school year.

Both classes require a passing score on the end-of-course assessment to graduate (Ruff, 2019). After citing permission and approval, a request was put through the district's data office and approved. The archived assessment data were pulled from the district's database, prepared by district employees, and emailed for the research study. There was no identifiable information located in the data file. The data file consisted of ELL SOL assessment scores from World History I and Algebra I during the 2022–2023 school year. A sample size of n-376 was drawn out of 2,958 World History I SOL scores and 2,380 Algebra SOL scores. There were no deviations from the required sample size due to the many test scores received.

Data Preparation

After the data file and research data were collected from the research site, the data for the quantitative causal comparative analysis was prepared in several ways. The data file was organized and prepared by the district's data office. A district data specialist pulled the school from the database and included the course code. This course code determined the ELL classroom placement when completing the Algebra I and World History I SOL end-of-course tests. The district removed all data points that were invalid scores or exemption codes and provided only valid ones. In other words, the district prepared and organized most of the data. Finally, the data file was sent and was organized based on the parameters of this study.

The data sample was organized using Microsoft 365 Excel to extract a random sample from the data set. First, all data were separated into Algebra I and World History I scores, and these scores were further separated into separate columns for cotaught and self-contained. All data scores were then assigned a random number through Microsoft 365 Excel, randomizing the data set. A random sample was collected: n-376 Algebra I and World History I test scores, respectively. The selected randomized data sets were placed in two different folders. One was for

Levene's tests, and the other was for the independent two-sample *t*-test. Lastly, the data sets were placed into different Microsoft 365 Excel files where statistical testing and analysis were completed.

Data Analysis and Results

This quantitative causal comparative study used *t*-tests to determine if there was a statistically significant difference in mean scores between two data sets (Middlemis Maher et al., 2013). According to DeMoulin and Kristonis (2009), *t*-tests are used to determine if the null hypothesis can be accepted or rejected and provide a level of confidence in that decision. The *t*-test in this study determined whether there was a significant mean difference between the two mean scores of cotaught and self-contained ELL assessment scores. When the difference in the mean scores increases, causing the *t-value* to increase and the *p-value* to decrease, the *p-value* depends on several assumptions to be considered accurate and valid (Suter, 2011). There are three main assumptions. The population from which the sample was drawn is normal and evenly distributed; the variance amount of the SOL scores is equal; and the sample size is randomly distributed (Suter, 2011). T-tests are most accurate when all assumptions are met. An alpha level of .05 was used in all decisions based on statistical significance.

Normality assumptions in this research were based on two factors. First, the central limit theorem (CLT) states that larger sample sizes mean a more excellent normal distribution of the data points (Sheposh, 2023). The study used a large sample of $n>50$, so any skewness had little effect on the results, as most research with $n>30$ is accepted (Sheposh, 2023). Secondly, a Q-Q scatterplot tests the assumptions of normality by plotting data points against the theoretical distribution. Considering most data points fall along the linear trendline, a large sample size and

regular data distribution in CLT can be assumed. Figures 1 and 2 present the scatterplot for the distribution of cotaught and self-contained ELL Algebra I and World History I SOL scores.

Figure 1

QQ Scatterplot of Normality in Algebra SOL Dependent on Classroom Placement

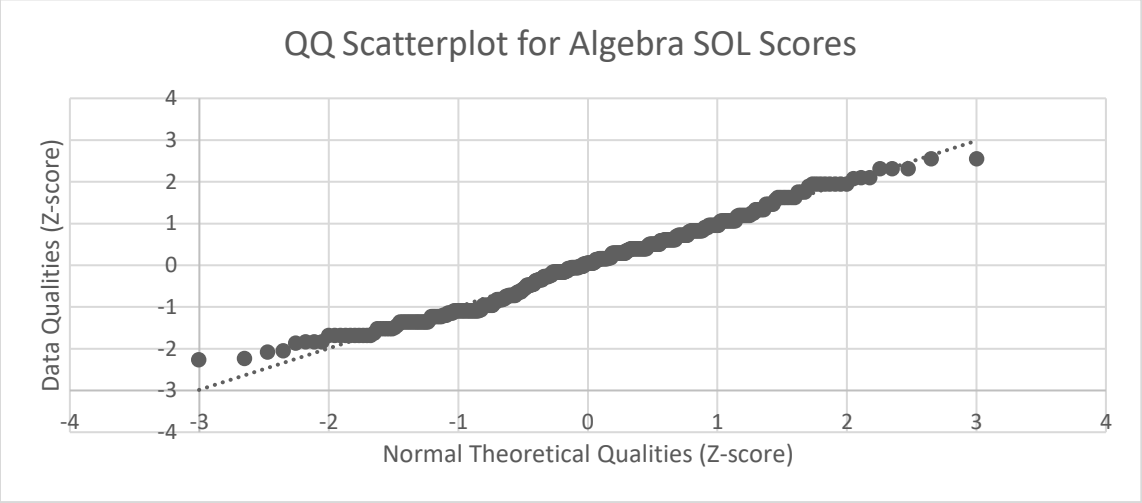
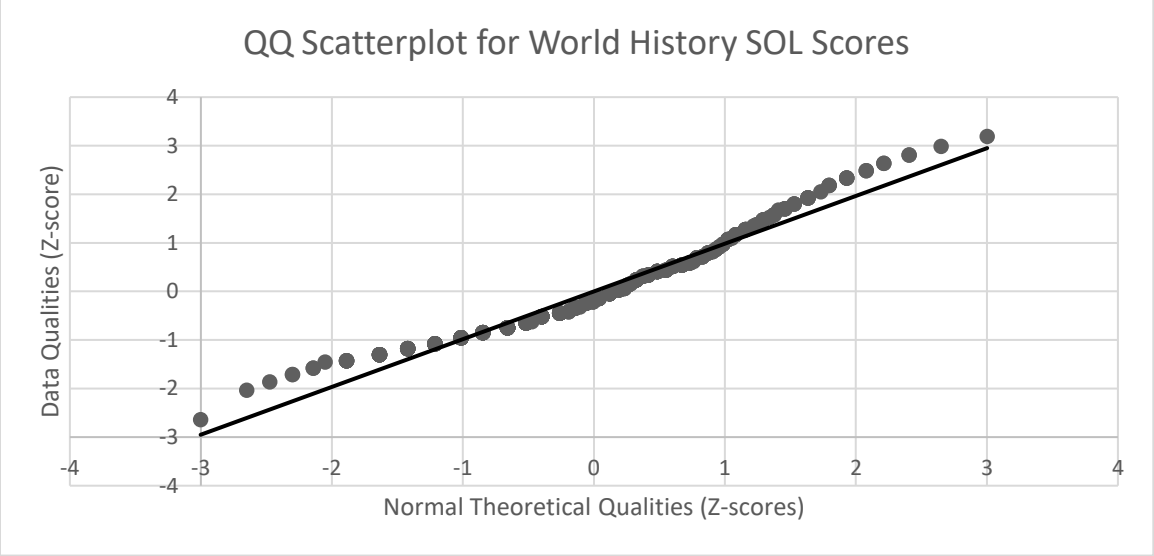


Figure 2

QQ Scatterplot of Normality in World History SOL Dependent Classroom Placement



Levene's test for variance equality was used to examine whether the homogeneity of variance assumption was met (Levene, 1960). Homogeneity is established when the variance of the dependent variable is equal among each group (Levene, 1960). The results of Levene's test on the Algebra I SOL data were $F = 0.06$, $p = .79$, and the World History I SOL scores were $F = 1.37$ and $p = .24$. Neither result was significant. This indicated that the assumption of homogeneity variance was met (Levene, 1960). Table 2 illustrates these results.

Table 2

The Results of Levene's Test on Each Data Set

Measure	F	df	P
Algebra SOL Scores	0.06	375	.79
World History SOL Scores	1.37	375	.24

An independent samples *t*-test was used because the SOL score meant that being tested was not connected and was a separate assessment. The research questions compared ELL self-contained assessment scores and cotaught independent scores with no connection. Both research questions were framed using different content-area assessments to compare ELL performance. The ELL classroom placement was different, cotaught, and self-contained, so an independent samples *t*-test was appropriate because all assumptions were also met. Each research question's data set met all assumptions; therefore, an independent two-sample *t*-test assuming equal variance was performed.

Microsoft 365 Excel can perform a variety of data analysis tools that are built into the program. A new Microsoft 365 Excel file was created, and the randomized sample for each research question was pasted into a separate sheet in Microsoft 365 Excel. The data analysis tool was used, and an independent two-sample *t*-test assuming equal variances was selected. The columns containing the first variable, cotaught ELL classes, were selected, and the second variable, self-contained ELL classes, was selected. Lastly, an output cell was chosen, and Microsoft 365 Excel performed the statistical procedures. This process was repeated for both research questions and data sets.

Research Question 1

The results of the first research question indicated a statistically significant difference in Algebra I SOL scores depending on ELL classroom placement. The independent *t*-test results were significant, t = 1.97, p = .05. The p-value met the critical threshold, and the null hypothesis could be rejected for the alternative (DeMoulin & Kritsonis, 2009). There was a significant difference in SOL scores based on classroom placement. The Algebra I SOL cotaught scores (M = 402) significantly differed from the Algebra I SOL self-contained scores (M = 409). A score of 400 is passing on these assessments. Table 3 illustrates these results.

Table 3

Results of Two Sample t-Test Assuming Equal Variances for Research Question 1

	Cotaught	*Self-Contained*
Mean	402.0053191	409.4840426
Variance	1399.128314	1404.657498
Observations	188	188
Pooled Variance	1401.892906	
Hypothesized Mean Difference	0	
Df	374	
t Stat	-1.936573475	
P(T<=t) one-tail	0.026775236	
t Critical one-tail	1.648938048	
P(T<=t) two-tail	0.053550472	
t Critical two-tail	1.966327183	

Research Question 2

The results of the second research question indicated a statistically significant difference in ELL World History SOL scores depending on their classroom placement. The results of the *t*-test were significant, t = 1.96, p = 0.51. The p-value exceeded the critical threshold, and the null hypothesis could also be rejected. There was a significant difference in World History SOL

scores depending on cotaught or self-contained classes. The statistical difference in the cotaught World History SOL mean scores (M = 384) and self-contained mean scores (M = 387). However, both classroom models still need to produce a passing mean score. Table 3 illustrates these results.

Table 4

Results of Two Sample t-Test Assuming Equal Variances for Research Question 2

	Cotaught	*Self-Contained*
Mean	384.4840426	387.2393617
Variance	1453.374076	1828.257908
Observations	188	188
Pooled Variance	1640.815992	
Hypothesized Mean Difference	0	
Df	374	
t Stat	-0.659486481	
P(T<=t) one-tail	0.254994648	
t Critical one-tail	1.648938048	
P(T<=t) two-tail	0.509989297	
t Critical two-tail	1.966327183	

A null hypothesis is rejected depending on the *p* level, which is the t-test outcome determining a statistical significance (Suter, 2011). Both statistical procedures produced *p*-values that met or exceeded the critical values set by the research parameters. The p-value of .05 indicates a significant difference in SOL scores between cotaught and self-contained classes with only a 5% margin of error (Mohn, 2022). In addition, all assumptions were met, giving the statistical results a high confidence level and improving reliability and validity (DeMoulin & Kritsonis, 2009).

Reliability and Validity

Reliability and validity of a research study refers to the study's ability to obtain valid results that can be generalized and replicated in other studies (George, 2021). Validity refers to the measure of the data point and reflects an accurate and consistent measure (Wienclaw, 2021). Reliability refers to a study's ability to produce the same results (Wienclaw, 2021). These components are critical aspects of all research and significantly affect findings and generalizability. This study's large, randomized sample addresses many threats to reliability and validity (George, 2021). However, additional steps were taken during this study to amplify the validity and reliability, thus amplifying the findings. The measures in each variable of this study were validated standardized tests scores used to assess students' content knowledge needed for graduation in the state of Virginia (Konold et al., 2018). This measure was used to increase the internal validity of this research study. According to the Virginia Department of Education, the SOL assessments are validated through a rigorous and precise process and have a reliability of >.80 overall (Konold et al., 2018). Therefore, this measure accurately measures students' understanding of academic content.

In addition to valid and reliable data, reliability and validity measures were performed to enhance the statistical analysis to ensure all statistical assumptions were met. A large, randomized sample size ($n>30$) meets two assumptions requirements to be valid and reliable (Sheposh, 2023). Meeting statistical assumptions and verifying that assumptions are met also strengthens the generalizability of this study (DeMoulin & Kritsonis, 2009). Statistical testing of data sets ensured normal distribution of data and homogeneity in the variance of SOL scores. Together, these safeguards sought to improve the external validity of this study and its generalizability to larger ELL populations.

Each research study must address various threats to validity and reliability that limit different aspects of each research study. This study measured ELL performance on a standardized test in their second language. There was no account for English language proficiency level in this study. The theoretical framework established that each ELL was placed in the class that best fits their needs. This research investigated a statistical difference in the ELL classroom placement meeting the needs of ELL populations by completing graduation requirements (Ruff, 2019). A large random sample and statistical assumption validation alleviated many threats posed by language proficiency levels.

A large sample size, valid and reliable measures, and statistical assumptions testing all work together to improve the validity and reliability of this research study. Further study could incorporate language proficiency levels and examine the data further. This research study revealed a significant difference, and further investigation is warranted. Nevertheless, this study has a large enough sample of participants and statistical validation measures that amplify the generalizability.

Chapter Summary

Quantitative comparative research studies can be used to determine if a statistical difference exists when comparing two different independent variables (Suter, 2011). This research examined a large sample (n-376) of ELLs in Northern Virginia in Grades 9–12. T-tests were used to examine a significant difference in test scores in high school ELLs depending on their classroom placement. In this research context, both null hypotheses were rejected due to the *p*-value meeting or exceeding the critical value. In addition, assumption testing ensures that each statistical procedure has little threat to validity and reliability.

The *t*-tests used for Research Question 1, which focused on Algebra SOL scores, produced the following results: $t = 1.97$, $p = .05$. These results were significant, and the null hypothesis was rejected for the alternative. Likewise, the results of Research Question 2 were $t = 1.96$ and $p = 0.51$, and they were also significant, causing the acceptance of the alternative hypothesis. These results showed a significant difference in test scores based on ELL classroom placement and could factor into their ability to meet graduation requirements (Konold et al., 2018).

A significant difference in both data sets indicated that ELL classroom placement is critical to addressing ELL achievement gaps in public schools (Voss, 2021). A large sample of test scores combined with homogeneity of variance and normal data distribution indicated strong statistical results with high confidence and a low margin of error. Therefore, the results of this statistical analysis will be used as the foundation for further discussion of the findings and implications of the research questions. In the final chapter, findings will be interpreted, results reported, and conclusions drawn. Recommendations for the future and implications for leadership are provided.

Chapter 5: Discussion and Conclusions

This quantitative comparative research study aimed to examine English language learners (ELLs) standardized assessment scores based on their classroom placement for any significant difference. Determining a significant difference involved gathering test scores (n=376) and organizing them based on ELL classroom placement to make this comparison. Communication and language input are needed for ELLs to develop competence in English and engage in discourse in the classroom. These interactions between ELLs and peers and ELLs with teachers strengthen content and language development. Using the classroom social interactions as two broad categories, a statistical analysis was conducted to address the following research questions regarding ELL assessment performance.

The first research question examined how ELL placement in cotaught, or self-contained classes affected their scores on the Algebra I Standards of Learning assessment. This study found a significant difference in student achievement depending on the classroom setting. Both classroom models scored above a passing score of 400. Scores of cotaught classes, with a mean score of 402, and self-contained, with a mean score of 409, indicate that ELLs have a better conceptual understanding of the Algebra I content. In addition, the Algebra I concepts and standards are more accessible to ELL populations in both classroom models. Furthermore, self-contained ELLs scored higher overall, indicating that the self-contained classroom models are more desirable in meeting high school graduation requirements.

The second research question investigated how ELL classroom placement affected their World History I Standards of Learning assessment scores in world history. It produced a significant difference between cotaught and self-contained assessment scores. Cotaught classrooms scored a mean of 384, and self-contained classrooms scored a mean of 387. Both

scores were below the proficiency standard of 400 set by the state of Virginia. However, this significant difference in mean scores indicated that classroom placement impacts ELL achievement as measured by the World History I assessment. Self-contained ELL classrooms performed better than cotaught classrooms in both data sets.

Based on these findings, instructional models impact how ELLs perform on standardized content assessments. In addition, content that is more linguistically demanding and requires more considerable cultural knowledge is less accessible in ELL populations, regardless of the classroom setting. These findings have implications beyond the scope of this study in educational, social, and language development research. The following sections elaborate on these findings, limitations, and researcher recommendations and provide implications for educational leadership.

Findings, Interpretations, and Conclusions

This study found that both research groups performed better on standardized assessments in the self-contained classroom model. In this classroom model, ELLs are grouped with other ELLs of the same language proficiency and with a teacher certified in the content and ESOL (Sugarman, 2018). In both research assessment measures, self-contained cohorts scored significantly above cotaught. These results imply that ELLs are more likely to understand content presented exclusively for them than instruction designed for native English speakers. School districts attempt to service ELLs by providing academic and language development to address achievement gaps (Goodrich et al., 2020). These findings also supported current literature on culturally relevant pedagogy and the importance of developing a curriculum that supports learners' experiences and building a supportive learning environment (Waxman et al., 2021). Interactions and student communication about content are fundamental aspects of

education and language pedagogy. In addition, examining these findings through sociocultural theory has implications for education and language development research. It also confirms research themes on the importance of professional development in preparing teachers to meet the educational needs of ELLs (Szymanski & Lynch, 2020).

Current educational research highlights the benefits of having a learner-centered environment that improves academic outcomes and boosts learner's self-efficacy (Banse & Palacios, 2018). The findings in this research supported these conclusions. The self-contained environment is designed to be a more sheltered form of instruction specifically for ELLs (Sugarman, 2018). At the same time, cotaught classes are designed for English-speaking students as the primary learners. On both assessments, students in the self-contained classroom setting performed significantly in vastly different content areas. These findings supported research validating improved learning outcomes by creating learning environments that support social and emotional needs (Kibler & Chapman, 2019). This indicated that smaller class sizes with increased language support for ELLs provide better student outcomes on standardized assessments required for graduation.

Schools use ELL classroom placement to match learners' communication and learning needs best while creating a supporting learning environment. Schools create these learning environments to maximize content understanding and language development. Educators use various supports and educational strategies to enhance the learning environment (De Jager, 2019). This study's results indicated these supports are an essential part of the educational process for ELLs. These findings supported a body of research providing evidence that supportive educational environments produce better student outcomes (Register et al., 2022). In addition, the learning environment determines how students communicate and participate in

collaborative learning (Snow, 2018). Collaborative learning is essential to learner engagement, and communication engages higher-order thinking skills and content understanding (Abulhassan & Hamid, 2021). Communication about content is more successful when ELLs are placed in learning environments specific to their language needs.

The findings of this quantitative causal comparative research have implications for education and language acquisition research. Sociocultural theory, proposed by Vygotsky in 1934, states that all learning requires social interactions. Learning new knowledge is best learned when content and language meet the learners' communication abilities (Vygotsky, 1978). In other words, educators are meeting the learners' current abilities, linguistically and academically, by placing ELLs in supportive learning environments and fostering communication about content. A significant difference in test scores indicated the importance of placing ELLs based on their current language level when learning challenging academic content.

Furthermore, social development theory refers to the optimal place of linguistic and cognitive development as the learner's zone of proximal (ZPD; Vygotsky, 1978). Schools and educators can meet the needs of ELLs by providing lessons that accommodate and communicate in ELLs' ZPD. This applies equally to each content area as self-contained ELLs performed higher in different content areas. Educators create the learning environment and differentiate instruction to accommodate learners' zone of proximal development, and these results confirmed these findings (Vintan & Gallagher, 2019). Examining these findings through this theoretical framework has implications for how ELLs learn language and content depending on the educational setting and communicative interactions.

Research literature focused on teachers' preparedness to meet the needs of ELLs in academic content classes emphasizes the importance of educator training. Qualitative data show

that educators report disparities in their abilities to meet ELLs' linguistic and academic needs (Szymanski & Lynch, 2020). Educators trained to support these needs report more confidence in meeting educational standards (Song, 2016). This study's significant difference in algebra assessment scores supported these findings because the team and self-contained teachers successfully supported ELLs' needs, as indicated by passing scores. However, world history team teachers and self-contained teachers did not produce passing results despite having a significant difference in score data. This finding suggested that the content is critical to ELLs' ability to access the curriculum and the benefits of educator training and development (Li & Peters, 2020). History content relies heavily on language use and cultural knowledge for students to access the curriculum. Teacher preparedness and development are paramount to learner success. Moreover, research on professional development for ELL educators can incorporate content-specific approaches.

The interpretations presented because of these findings are within the scope of this research study for several reasons. In Algebra I cotaught and self-contained classes, ELLs scored above the required state scores of 400, indicating that each classroom setting correctly supports ELL achievement. These findings also indicated that algebra content is more accessible to ELLs regardless of classroom placement. World History I content is not as accessible for ELLs despite a significant score difference. Both mean scores were below the passing threshold. Using content assessments allowed research conclusions to be limited without consideration of their English proficiency levels.

Limitations

This quantitative causal comparative study used a statistical analysis of ELLs' performance on standardized tests in two academic content areas. The data set comprised a large,

randomized sample (n=376) out of several thousand assessment scores. A large sample size, pre-vetted assessment data point, and adherence to all statistical assumptions improved this research study's internal and external validity.

The data collected examined students' performance on standardized assessments needed to graduate high school in Virginia. These results shape policy and influence educational decisions at state and local governmental levels (Ruff, 2019). Virginia Standards of Learning (SOL) assessment is a valid and reliable measure of students' achievement in several ways. First, these assessments are criterion-referenced assessments developed and verified through independent assessment specialists (Virginia Department of Education, 1999). These assessments and each teaching item are subject to continued vetting and review by independent testing organizations and have a reliability of >.80 (Konold et al., 2018). Therefore, these assessments are considered a valid and reliable measure, and a large sample size strengthens these findings' internal and external validity.

Moreover, this study was limited by only measuring ELL academic performance for a significant difference depending on their instructional settings through a communicative theoretical framework. This study did not account for ELLs' language proficiency level. However, generally, school districts tend to place lower-level ELLs in self-contained classrooms to offer more support in language development and content instruction (De Jager, 2019). The results indicated a significant difference in assessment data, which supported research related to ELLs' educational environment and the application of sociocultural theory. Regardless of English proficiency level, placing ELLs based on their level of communicative competence improves educational outcomes for these students. Lastly, ELL assessment data as the only point of measurement omitted many other variables that affect educational outcomes in ELL

populations. For example, larger districts have more educational training, support, and technology resources. In addition to these variables, educators all present content in various ways and may incorporate students' native language within instruction. Variables of this nature can damage the reliability of this study; however, the strict application of statistical analysis mitigated these limitations.

A large, randomized sample (n=376) strengthened the statistical analysis results. All statistical assumptions were met before statistical analysis, further strengthening the results and their implications. A valid and reliable assessment measure was used to measure student performance in these subject areas. Internal and external validity measures produce findings generalizable to larger ELL populations. These findings contribute to the literature surrounding education and language development, and further recommendations can be made based on these findings.

Recommendations

Creating positive and desirable student learning outcomes is a complex and multifaceted process affected by countless variables. Research-based practices aim to improve educational outcomes and the learning experiences of all students. However, the educational landscape continues to evolve and change, leading to constant revisions in how educational services are delivered. Therefore, educational research should evolve with this ever-changing landscape. This study indicated a significant difference in ELL academic performance depending on the learning environment. Further research should continue to investigate the difference by measuring language acquisition levels, language arts standardized assessment scores, and various qualitative measures.

Measure Annual Language Acquisition Assessments

The results of this study indicated that ELL classroom placement can yield a significant difference in academic achievement. Researchers should examine annual testing data measuring ELL language proficiency development. School districts should assess ELLs yearly to determine their language proficiency level for scheduling and resource management (Sugarman, 2018). This study found that classroom setting significantly impacts ELL academic performance on content assessments. Statistical analysis using language assessment scores as the dependent variable could determine which classroom model improves language acquisition.

Examine 11th-Grade English Standards of Learning Assessment

This study focused on two subject areas that use language differently to help ELLs access and understand the content. In Virginia, high school students are expected to pass an English content assessment focusing on reading and writing (Virginia Department of Education, 1999). Researchers should perform a similar statistical analysis of this language arts assessment. This assessment requires students to read passages and answer questions, and grammar and writing. This assessment is particularly challenging for ELLs due to the complexity of the language involved, and alternative assessments are often given to satisfy this graduation requirement (Ruff, 2019).

Identify the Best Practices for Each Class Model

A consistent theme in surveyed educators regarding ELLs is that teachers expect professional development to give them proven effective strategies (Song, 2016). Quantitative assessment can only identify a significant difference in test scores and needs to identify standard variables contributing to student achievement. Researchers should survey students and teachers to determine which best practices improve student outcomes. This can help institutions

determine which educational practices work best for their students. In addition, public school districts can further support ELL educational needs by placing learners in self-contained classrooms that require standardized assessments.

Implications for Leadership

The quantitative causal comparative study has implications for educational leadership that involve improving ELL achievement in required content classes. These findings suggest that educational leadership can improve ELLs' achievement in several ways. ELL learning environments can potentially improve educational experiences and close achievement gaps, and this study supports this conclusion (Voss, 2021). Self-contained ELL educators are dual certified, and this study confirms that their training produces higher academic achievement regardless of language proficiency. Finally, educational leaders can make administrative decisions, like student placement, to produce better educational outcomes within high-needs populations.

Create Supportive Learning Environments

The learning environment is an essential part of the educational process, from the teacher and curriculum to the physical layout of the classroom and school. Educational leaders can support ELL students by creating supporting learning environments with trained professionals who are more equipped to meet ELL needs. This sentiment is supported by educational research and language acquisition (Dos Santos, 2020). The findings of this study add to the growing body of research highlighting the benefits of a supportive learning environment that accommodates different learning styles and students. These findings also extended to secondary education based on this study's results. Educational leaders can create programs and select educators to create a supportive learning environment for ELL success.

Teacher Training for ELL Support

The statistical analysis found a significant difference in assessment data depending on classroom placement, which highlights the importance of teacher training. Self-contained teachers are certified in content and have taken an assessment for an ESOL endorsement (Sugarman, 2018). These results indicated that continued professional development improves student outcomes. Educational leaders should create a continuous educator growth and development culture within educational institutions (Fisher & Carlyon, 2015s). Educators with a wide range of skillsets are more capable of supporting ELL academic development than ESOL and content teachers. Educational leaders can consider this while hiring teachers and fiscal resource management.

Improving ELL Policy to Better Support Learners' Needs

Achievement gaps between ELLs and native English speakers continue to widen yearly despite ELLs receiving language development services and content support (Goodrich et al., 2021). Educational leaders can schedule and allocate resources to make content more accessible to ELL populations. Also, resource management can be concentrated in subjects that require more support for ELLs, such as language and cultural heavy content. This study indicates that more language and cultural-intensive subjects require more support for ELLs. Therefore, educational leaders can support ELLs through class scheduling focusing on self-contained classes that require high-stakes testing measures to complete.

Conclusion

In conclusion, self-contained ELL content classes produce higher scores on state standardized assessments required for graduation in Virginia. The findings from this research show the importance of appropriate classroom placement for ELL success. These findings can

support ELL educational experiences in Virginia, and the large, randomized sample size makes these results generalizable. These measures add to the body of literature surrounding education and language acquisition. In addition, these findings have implications in these same areas. Overall, the outcomes of this study suggest that educational institutions can support ELL academic development by placing them in learning environments that support their academic development.

This quantitative causal comparative study adds to the body of educational literature that emphasizes the need to support ELLs in academic development instead of just incorporating ELLs in general education classrooms. Even with two teachers, cotaught classrooms failed to meet the academic needs of ELLs. Classroom environment, teacher preparation, and curriculum content provide better opportunities for ELLs to engage with the academic curriculum. This study found that supporting these needs leads to better academic outcomes for ELLs.

The findings of this research have implications for language acquisition instruction and applicational educational research. First, further research using the same methodology would identify the importance of classroom placement and which content areas require the most significant allocation of resources. Factoring for language acquisition could allow educational institutions to serve ELL populations better and assimilate them into American schools. Lastly, educational research and practice are improved because educators can adjust content instruction to accommodate ELLs better, regardless of language proficiency.

References

Abulhassan, A. B. A., & Hamid, F. I. E. (2021). Perception and interest of English language learners (ELL) toward collaborative teaching: Evaluation towards group activities. *English Language Teaching*, 14(5), 1–12.

Adem, H., & Berkessa, M. (2022). A case study of EFL teachers' practice of teaching speaking skills vis-à-vis the principles of communicative language teaching (CLT). *Cogent Education*, *9*(1), 1–23. https://doi.org/10.1080/2331186X.2022.2087458

Artigliere, M. (2019). The proficiency, instructional and affective domains of long term English language learners: A review of the research. *TESL-EJ*, *23*(1), 1–19.

Banse, H., & Palacios, N. (2018). Supportive classrooms for Latino English language learners: Grit, ELL status, and the classroom context. *Journal of Educational Research*, *111*(6), 645–656. https://doi.org/10.1080/00220671.2017.1389682

Bardack, S. (2010). *Common ELL terms and definitions*. American Institutes for Research.http://www.air.org/sites/default/files/downloads/report/NEW-Common_ELL_TERMS_AND_DEFINITIONS_6_22_10_0.pdf

Besterman, K., Williams Jr., T. O., & Ernst, J. V. (2018). STEM teachers' preparedness for English language learners. *Journal of STEM Education: Innovations & Research*, *19*(3), 33–39.

Bloor, T., & Santini, J. (2023). Modeling the epistemic value of classroom practice in the investigation of effective learning. *Science & Education*, *32*(1), 169–197. https://doi.org/10.1007/s11191-021-00298-9

Bond, K. M. (2020). Two viewpoints of English learner (re)classification: The implications of federal requirements for defining English learners in U.S. schools. *TESOL Journal*, *11*(2), 1–5. https://doi.org/10.1002/tesj.491

Britsch, S., & Shepardson, D. P. (2021). The use of teacher resources in a sheltered science unit on watersheds. *Science Educator*, *28*(1), 41–53.

Campbell, Y. C., & Filimon, C. (2018). Supporting the argumentative writing of students in linguistically diverse classrooms: An action research study. *Research in Middle Level Education Online*, *41*(1), 1–10. https://doi.org/10.1080/19404476.2017.1402408

Case, A. F. (2015). Beyond the language barrier: Opening spaces for ELL/non-ELL interaction. *Research in the Teaching of English*, *49*(4), 361–382. https://go.openathens.net/redirector/ace.edu/login?url=https://www.proquest.com/scholarly-journals/beyond-language-barrier-opening-spaces-ell-non/docview/1683507443/se-2

Castrillón, L. J. V. (2017). The effects of Vygotsky's sociocultural theory on second language acquisition and language Input. *ESPIRAL: revista de docencia e investigación*, 7(1), 91–102. https://doi.org/10.15332/erdi.v7i1.1780

Çiftçi ertürk, D., & Özdemir, M. (2022). Analysis of the effects of factors, which affect language acquisition, on four basic language skills. *Sakarya University Journal of Education*, *12*(1), 206–223. https://doi.org/10.19126/suje.1058010

Clark, C. P. (2021). An opportunity for change: Groundbreaking scholar Gloria Ladson-Billings on culturally relevant pedagogy and why education as we know it needs to be transformed. *Literacy Today*, *38*(5), 24.

Cole, M. (2018). Effectiveness of peer-mediated learning for English language learners: A meta-analysis. *Research Ideas and Outcomes*, *4*(1–77), 1–77. https://doi.org/10.3897/rio.4.e29375

Creswell, J. W. (2009). *Research design: Qualitative, quantitative, and mixed methods approaches* (3rd ed.). Sage Publications.

Dan Guo. (2018). A study of English service learning under the theory of second language acquisition. *International Forum of Teaching & Studies*, *14*(2), 27–33.

De Jager, T. (2019). Are principles of the Sheltered Instruction Observation Protocol Model promoting ESL teaching and learning? *International Journal of Pedagogy & Curriculum*, *26*(1), 43–58. https://doi.org/10.18848/2327-7963/CGP/v26i01/43-58

DeMoulin, D. F., & Kritsonis, W. A. (2009). Sampling and hypothesis testing. *In a Statistical Journey: Taming of the Skew*.

Derwing, T. M., Decorby, E., Ichikawa, J., & Jamiesen, K. (1999). Some factors that affect the success of ESL high school students. *Canadian Modern Language Review*, *55*(4), 532–547.

Desjardins, A. (2020). Effective sheltered content instruction for English language learners. *BU Journal of Graduate Studies in Education*, *12*(2), 15–18.

Dhai, A. (2017). Exploitation of the vulnerable in research: Responses to lessons learnt in history. *South African Medical Journal*, *107*(6), 472–474. https://doi.org/10.7196/SAMJ.2017.v107i6.12437

Dos Santos, L. M. (2020). The discussion of communicative language teaching approach in language classrooms. *Journal of Education and E-Learning Research*, *7*(2), 104–109.

Eichhorn, M. S., Lowry, A. E., & Burke, K. (2019). Increasing engagement of English learners through Universal Design for Learning. *Journal of Educational Research & Practice*, *9*(1), 1–10. https://doi.org/10.5590/JERAP.2019.09.1.01

Fisher, A., & Carlyon, T. (2015). School leaders growing leadership from within: A framework for the development of school leaders. *Waikato Journal of Education (2382–0373)*, *20*(2), 93–102. https://doi.org/10.15663/wje.v20i2.179

Fraenkel, J. R., & Wallen, N. E. (2012). *How to design and evaluate research in education* (7th ed.). McGraw-Hill.

George, A. C. (2021). A brief guide to sampling in educational settings. *Tutorials in Quantitative Methods for Psychology*, *17*(3), 286–298. https://doi.org/10.20982/tqmp.17.3.p286

Goodrich, J. M., Thayer, L., & Leiva, S. (2021). Evaluating achievement gaps between monolingual and multilingual students. *Educational Researcher*, *50*(7), 429–441. https://doi.org/10.3102/0013189X21999043

Guillén, G., Sawin, T., & Avineri, N. (2020). Zooming out of the crisis: Language and human collaboration. *Foreign Language Annals*, *53*(2), 320–328. https://doi.org/10.1111/flan.12459

Guy, B., Shin, H., Lee, S., & Thurlow, M. (2000). State graduation requirements for students with and without disabilities. Issues influencing the future of transition programs and services in the United States, 85–110.

Hackett, J., Kruzich, J., Goulter, A., & Battista, M. (2021). Tearing down the invisible walls: Designing, implementing, and theorizing psychologically safer co-teaching for inclusion. *Journal of Educational Change*, *22*(1), 103–130. https://doi.org/10.1007/s10833-020-09401-3

Hahl, K., & Pietarila, M. (2021). Class teachers, subject teachers and double qualified: conceptions of teachers' skills in early foreign language learning in Finland. *International Electronic Journal of Elementary Education*, *13*(5), 713–725. https://doi.org/10.26822/iejee.2021.223

Howard, T. (2013). *Why race and culture matter in our schools: Closing the achievement gap in America's classrooms*. Teachers College Press.

Howlett, K. M., & Penner-Williams, J. (2020). Exploring teachers' perceptions of an English language proficiency (ELP) standards professional development workshop. *TESL-EJ*, *24*(2), 1–20.

Kataike, J., Kulaba, J., & Gellynck, X. (2017). What junior researchers must know before and after data collection: Difference between parametric and nonparametric statistics.

Kibler, K., & Chapman, L. A. (2019). Six tips for using culturally relevant texts in diverse classrooms. *Reading Teacher*, *72*(6), 741–744. https://doi.org/10.1002/trtr.1775

Kieffer, M. J., & Thompson, K. D. (2018). Hidden progress of multilingual students on NAEP. *Educational Researcher*, *47*(6), 391–398.

Konold, T., Cornell, D., Jia, Y., & Malone, M. (2018). School climate, student engagement, and academic achievement: A latent variable, multilevel multi-informant examination. *Aera Open*, *4*(4), 2332858418815661.

Krashen, S. D. (1982). *Principles and practice in Second language acquisition*. Pergamon Press.

Lailah Isrofatun Nahdiah. (2022). Teacher and students' perceptions toward the use of students' L1 in Efl Classroom. *Religio Education*, *2*(1), 11–20. https://doi.org/10.17509/re.v2i1.46764

Lertcharoenwanich, P. (2022). The effect of communicative language teaching in test preparation course on TOEIC score of EFL business English students. *Journal of Language Teaching & Research*, *13*(6), 1188–1195. https://doi.org/10.17507/jltr.1306.06

Levene, H. (1960). *Contributions to probability and statistics. Essays in honor of Harold Hotelling*, 278–292. Stanford University Press.

Li, C., Kruger, L. J., Beneville, M., Kimble, E., & Krishnan, K. (2018). The unintended consequences of high-stakes testing on English-language learners: Implications for the practice of school psychology. *School Psychology Forum*, *12*(3), 79–90.

Li, N., & Peters, A. W. (2020). Preparing K-12 teachers for ELLs: Improving teachers' L2 knowledge and strategies through innovative professional development. *Urban Education*, *55*(10), 1489–1506.

Liping W. (2021). Teaching academic vocabulary to English language learners (ELLs). *Theory & Practice in Language Studies (TPLS)*, *11*(12), 1507–1514. https://doi.org/10.17507/tpls.1112.01

Mayer, A., LeChasseur, K., & Donaldson, M. (2018). The structure of tracking: Instructional practices of teachers leading low- and high-track classes. *American Journal of Education*, *124*(4), 445–477. https://doi.org/10.1086/698453

Meskill, C., & Oliveira, A. W. (2019). Meeting the challenges of English learners by pairing science and language educators. *Research in Science Education*, *49*(4), 1025–1040. https://doi.org/10.1007/s11165-019-9837-9

Middlemis Maher, J., Markey, J. & Ebert-May, D. (2013). The other half of the story: Effect size analysis in quantitative research. *CBE-Life Science Education* (12), 345–351.

Mohn, E. (2022). p-value. *Salem Press Encyclopedia of Science*.

Novitasari, A., Raja, P., & Flora, F. (2018). Designing collaborative blended learning activities to improve students' argumentative essay writing ability and students' perception. *U-JET, 7*(4). https://doi.org/10.23960/jpp

Olsen, L. (2014). *Meeting the unique needs of long term English language learners*: *A guide for educators.* National Education Association.

Owens, C. W., & Wells, S. P. (2021). Elementary content teacher perceptions regarding their ELL instructional practices. *Journal of Educational Research & Practice, 11*(1), 139–152.

Payant, C., & Bell, P. (2022). Developing literacy skills through collaborative tasks for emerging-proficiency English as additional language learners in Quebec. *TESL Canada Journal*, *39*(1), 1–15. https://doi.org/10.18806/tesl.v39i1/1369

Pumpki Lei Su, Rojas, R., & Iglesias, A. (2022). Dual language profiles in Spanish-speaking English learners. *Journal of Speech, Language & Hearing Research*, *65*(7), 2608–2628. https://doi.org/10.1044/2022_JSLHR-21-00447

Register, J., Fernandes, A., & Pugalee, D. (2022). Supporting preservice mathematics teachers' culturally responsive teaching: A focus on teaching for social justice. *Mathematics (2227– 7390), 10*(6), 896. https://doi.org/10.3390/math10060896

Rizzuto, C. K. (2017). Teachers' perceptions of ELL students: Do their attitudes shape their instruction? *Teacher Educator, 52*(3), 182–202.

Ruff, R. R. (2019). State-level autonomy in the era of accountability: A comparative analysis of Virginia and Nebraska education policy through No Child Left Behind. *Education Policy Analysis Archives, 27*, 6. https://doi.org/10.14507/epaa.27.4013

Samson, J., & Collins, B. (2012). *Preparing all teachers to meet the needs of English language learners: Applying research to policy and practice for teacher effectiveness*. Center for American Progress. http://files.eric.ed.gov/fulltext/ED535608.pdf

Schenker, J. D., & Rumrill, J. P. D. (2004). Causal-comparative research designs. *Journal of Vocational Rehabilitation, 21*(3), 117–121.

Shen, X., & Byfield, L. (2018). Promoting English learners' willingness to communicate in content-area classrooms. *Clearing House, 91*(6), 250–257. https://doi.org/10.1080/00098655.2018.1541856

Sheposh, R. (2023). Central limit theorem. *Salem Press Encyclopedia of Science*.

Simon, M. K., & Goes, J. (2013). *Scope, limitations, and delimitations*.

Snow, C. E. (2014). Extended discourse in first and second language acquisition: A challenge and an opportunity. *Grantee Submission, 30*, 7–14.

Song, K. H. (2016). Systematic professional development training and its impact on teachers' attitudes toward ELLs: SIOP and guided coaching. *TESOL Journal, 7*(4), 767–799.

Spickard, J. (2017). *Step 6: pick a data analysis method*. SAGE Publications, Inc. https://dx.doi.org/10.4135/9781071802731

Sugarman, J. (2018). *A matter of design: English learner program models in K–12 education*. Issue Brief No. 2. Migration Policy Institute. 1–19.

Suresh, S. (2015). *Nursing research and statistics*. Elsevier Health Sciences. Suter, W. N. (2011).

Suter, W. N. (2011). *Introduction to educational research: A critical thinking approach* (2nd. ed.). SAGE Publications, Inc.

Swanson, H. L., Kong, J. E., Moran, A. S., & Orosco, M. J. (2019). Paraphrasing interventions and problem-solving accuracy: Do generative procedures help English language learners with math difficulties? *Learning Disabilities Research & Practice (Wiley-Blackwell), 34*(2), 68–84. https://doi.org/10.1111/ldrp.12194

Szymanski, A., & Lynch, M. (2020). Educator perceptions of English language learners. *Journal of Advanced Academics*, *31*(4), 436–450.

Taherdoost, H. (2022). What are different research approaches? Comprehensive review of qualitative, quantitative, and mixed method research, their applications, types, and limitations (August 1, 2022). *Journal of Management Science & Engineering Research*, *5*(1), 53–63, 2022. https://doi.org/10.30564/jmser.v5i1.4538

Ungvarsky, J. (2020). *Social development theory (Vygotsky).* Salem Press Encyclopedia.

United States Department of Education. (2017). Consolidated State Performance Report: Parts I and II. https://www2.ed.gov/admins/lead/account/consolidated/sy15-16part1/va.pdf

U.S. Department of Health & Human Services. (1979, April 18). *The Belmont Report*. https://iwww.hhs.gov/ohrp/regulations-and-policy/belmont-report/read-thebelmont-report/index.html

Vintan, A., & Gallagher, T. L. (2019). Collaboration to support ESL education: Complexities of the integrated model. *TESL Canada Journal*, *36*(2), 68–90. https://doi.org/10.18806/tesl.v36i2.1314

Virginia Department of Education. (1999). Virginia standards of learning assessments: Standards of Learning (SOL) tests validity and reliability information, spring 1998 administration.

Voss, L. (2021). Lost in translation? *The impact of increasing funding on high school graduation rates of English language learners*. UC Santa Barbara: Department of Economics. https://escholarship.org/uc/item/616845c4

Vygotsky, L.S. (1978). *Mind in society*. Harvard University Press

Vygotsky, L.S. (1934). *Thinking and speech*. The M.I.T. Press

Waxman, H. C., Padrón, Y. N., & Keese, J. (2021). Learning environment and students' classroom behavior differences between effective, average, and ineffective urban elementary schools for Hispanic students. *Educational Research for Policy & Practice*, 20(3), 307–324. https://doi.org/10.1007/s10671-020-09281-7

Wright, W. E. (2016). Let them TALK! *Educational Leadership*, *73*(5), 24–29.

Wienclaw, R. A. (2021). Reliability. *Salem Press Encyclopedia*.

Wienclaw, R. A. (2021). Validity. *Salem Press Encyclopedia*.

Yujuan Shi, Zaier, A., & Maina, F. (2021). Sheltered Instruction Observation Protocol Model: An effective way of promoting teacher candidates' self-efficacy in teaching English language learners. *International Journal of Diversity in Education*, *21*(2), 1–17. https://doi.org/10.18848/2327-0020/CGP/v21i02/1-17

Appendix A

Site Permission

[redacted]

August 21, 2023

Mr. W. Aaron Zimkowski

[redacted]

Dear Mr. Zimkowski,

The purpose of this letter is to inform you that your request to conduct doctoral level research in [redacted] titled "*Casual Quantitative Comparison of Cotaught and Self-contained High School ELL Classes for Differences in State Standardized Test Scores*" has been reviewed and approved by [redacted] leadership. Please coordinate research activities with staff in the Research, Accountability, and Strategic Planning Department.

Participation in the study is completely voluntary. Consent must be sought in all cases. In reporting the results please ensure anonymity of participants by removing all identifying information related to [redacted] and its staff. We look forward to reading your final results.

Thank you for your interest in [redacted] as a research site, and we wish you success!

Sincerely,

[redacted]

Director

Research, Accountability, and Strategic Planning