Fall 10-15-2021

Measuring Self-Efficacy of Elementary General Education Teachers Serving English Language Learners with and without Disabilities

Stacey L. McCrath-Smith
Seattle Pacific University

Measuring Self-Efficacy of Elementary General Education Teachers Serving English Language Learners with and without Disabilities

by

Stacey McCrath-Smith

Dissertation

Presented to the Faculty of the

Graduate School of Education at

Seattle Pacific University

In Partial Fulfillment of the Requirements for the

Doctor of Education Degree

Seattle Pacific University

2021

Measuring Self-Efficacy of Elementary General Education Teachers Serving English Language Learners with and without Disabilities

by

Stacey McCrath-Smith

A dissertation submitted in partial fulfillment

Of the requirements to the degree of

Doctor of Education

Seattle Pacific University

2021

Approved by: Cher N. Edwards
(Dr. Cher Edwards, Chairperson of the Dissertation Committee)

(Dr. Nyaradzo Mvududu)

Jorge A. Preciado
(Dr. Jorge Preciado)

Program Authorized to Offer Degree:

School of Education

Date: November 2021

(Dr. Nyaradzo Mvududu, Dean, School of Education)

Copyright Page

Dedication

I dedicate this accomplishment to the most important people in my life, my sons.

To Keith Charles Smith, you are my inspiration. You taught me, by your example, determination and discipline to keep going even in the face of monumental challenges.

To Stephen Reilly Smith, you are my motivation. You demonstrate and unwavering joy and persistence to keep moving forward, to embrace the unknown to achieve your dreams.

I love you both to the moon and back!

Mom

Acknowledgements

The process of undertaking the development of a study and writing a dissertation is a daunting task for any person who chooses that path, and for an individual, like me, who has attention deficit hyperactivity disorder (ADHD), it can present a nearly impossible challenge. However, prayer, persistence, determination, a great deal of deep breathing, refocusing and God's provision, saw me through the distraction, on a longer journey than most to the destination, a completed dissertation.

I would first like to thank the chairperson of my committee, Dr. Cher Edwards, for her guidance, patience and unfailing encouragement throughout the process. To Dr. Nyaradzo Mvududu for her positivity and direction through my analysis, and Dr. Jorge Preciado for his support and encouragement in development of my vision.

A thank you to my McCrath, Smith, Stiles family for your love and encouragement, Big Keith, Keith Charles, Stevie, and Keely, Mom and Bob, Dad, Steve, Katie, Kieran, Sienna and Kylian.

To my friends, Nicole Counts and Shana Valencia, I am so grateful for your love, encouragement, gentle nudging and steadfast belief in me, thank you for being my greatest cheerleaders.

And for getting me through the last stretch to the finish line, Dr. Whitney Meissner, Jill Vanderveer for my "quite place to focus", my "Boss Ladies" Debbie, Michelle and Kim, thank you for being in my corner and cheering me on, and especially Dr. Shannon Hitch, for providing me a model and motivation to encourage me every day.

"*Success is not final; failure is not fatal: It is the courage to continue that counts.*"

-Winston Churchill

Table of Contents

List of Figures

List of Tables

List of Appendices

Abstract

The K-12 student population across the United States has experienced a growth of students whose native language is not English (Piazza et al., 2020). These culturally, linguistically diverse students, also known as English Learners (ELs), are placed in classrooms of teachers who lack the training necessary to effectively serve language and content learning needs (Ortiz & Robertson, 2018). A model that has been routinely used to provide teachers with an instructional framework to address the needs of ELs, is the sheltered instruction observation protocol (SIOP) (Echevarria, Vogt, & Short, 2017). This study sought to develop a valid and reliable tool, the Sheltered Instruction Teacher Self-Efficacy Scale (SITSES) based on the components of the SIOP framework to measure the self-efficacy of general education classroom teachers serving ELs and to examine difference in self-efficacy ratings of teachers who had received SIOP training and those who did not. A pilot study was conducted to calibrate the tool and establish validity and reliability. Subsequently the tool was sent to 460 general education teachers in a public school district in the northwestern part of the United States resulting in a 10% response rate. Results of an exploratory factor analysis suggest that four components emerged, aligning with the eight SIOP instructional framework elements of lesson preparation, building background, strategies, interaction, practice/application, lesson delivery, review and assessment. In examination of the teacher self-efficacy scores, analysis suggests that teachers who have been trained in SIOP report a higher degree of efficacy in serving culturally, linguistically diverse learners than teachers who were not SIOP trained. Recommendations are made for future research and use of the tool to assess teacher self-efficacy serving ELs.

Chapter 1

Introduction

Historical Context

Over the past three decades, U.S. schools have seen a dramatic growth in the enrollment of bilingual students who often experience poor outcomes and face barriers in successfully accessing their education (Harris & Sullivan, 2017). America has received an influx of people from diverse cultural and linguistic backgrounds, many of them seeking relief from international conflicts, political oppression, failing economies, drought, and diminishing food supplies (Cummins, 1981; Piazza et al., 2020). Data collected by the National Center for Education Statistics (NCES) showed an increase in students who spoke a language other than English at home, from 8.1 % or 3.8 million students in 2008 to 10.2% or 5.0 million students in 2018 (NCES, 2021). Students learning English as a second language represent the fastest growing segment of the overall population (U.S. Census Bureau, 2010). In the fall of 2015, the percentage of English Learners (ELs) in American public schools was more than 10% in eight states (Alaska, California, Colorado, Kansas, Nevada, New Mexico, and Texas) with the highest percentage of 21% in California (NCES, 2018). While a large majority of ELs speak Spanish as a first language, ELs nationwide speak over 450 native languages. Demographically, these ELs vary in terms of socio-economic backgrounds, immigration status, schooling experience, and educational needs (Tran, 2011).

Roughly 7.6% of ELs in the United States have been identified as having a disability, with 55% of those identified as Learning Disabled (LD) (NCES, 2010). While many of these students are truly impacted by a disability, a significant number of ELs have been misidentified as having a disability due to the challenges they face in mastering academic content while at the

same time learning English (Sanchez et al., 2010). In 2006, Congress revised the Individuals with Disabilities Education Improvement Act (IDEIA; U.S. Congress, 2007) to address the over-identification of ELs requiring special education. Namely, multi-disciplinary teams considering a special education referral must establish that the difficulties experienced by culturally and linguistically diverse learners result from issues other than acquisition of a second language (L2) (U.S. Congress, 2007). Previous studies have identified challenges faced by practitioners in distinguishing between emergent English proficiency and learning disability (Keller-Allen, 2006).

Changing demographics in classrooms across the United States have placed a greater level of responsibility on general education teachers to expand the focus of their instructional practice from the delivery of content-based instruction implicitly geared toward native English speakers to a more inclusive style of teaching, one that incorporates an understanding of L2 development and of the differentiation in levels of instruction needed to reach all of their students (Nguyen, 2012; Ortiz & Robertson, 2018).

This change begs the question of how prepared teachers feel that they can adapt their instructional style to effectively teach culturally and linguistically diverse populations of ELs. As part of this work, it is also critical to understand what attitudes teachers hold and what knowledge and skills they believe they have obtained, as both of these elements contribute to how teachers perceive their teaching self-efficacy (Loreman, Sharma, & Forlin, 2013).

Given the challenges in effectively serving ELs, questions arise as to teachers of ELs efficacy with providing instruction to culturally and linguistically diverse learners (Ford, Stuart & Vakil 2014). Teacher efficacy is defined in the seminal work by Bandura (1977, 1995) as a teacher's belief in their capability to organize and execute courses of action to successfully

accomplish specific instructional tasks, or in other words, their capacity to affect student performances.

Problem Statement

Several studies have examined the pattern of ELs being identified and referred disproportionally as requiring special education services (DeValenzuela et al., 2006; Keller-Allen, 2006; Klingner et al., 2005; Maki et al., 2017; Sullivan, 2011). The U.S. Department of Education reported that in the fall of 2015 there were 713,000 ELs that were identified as students with disabilities. This number represented 14.7% of the total English learner population enrolled in U.S. public K-12 schools (NCES, 2018). Most referrals for a special education are made by general education classroom teachers, thus affirming that general education teachers require a better understanding of both second language development and learning disabilities than they currently possess (Ysseldyke, 2005). General education teachers benefit from increased instructional strategies designed to support ELs while simultaneously presenting subject matter content (Barker & Grassi, 2011). These deficits leave general education teachers ill prepared to serve ELs and those with disabilities, thus bringing into question their perceived self-efficacy. Additionally, ELs misidentified as LD are taught by special education teachers who lack the training and curriculum to meet students' needs in L2 and literacy development (Barker & Grassi, 2011; Zehler et al., 2003). Conversely, EL teachers well-versed in providing instruction to support the development of English as a second language lack the strategies to adequately address the needs of students with learning disabilities (Paneque & Barbetta, 2006)

The IDEA (2006) reauthorization addressed the overrepresentation of culturally and linguistically diverse students in referrals for special education services, by adding increased responsibility on general education teachers to meet the needs of English learner students. IDEA

called for the development of more supportive interventions to serve struggling students in the general education setting (Kaplan & Leckie, 2009). In the past decade, approaches such as GLAD (Guided Language Acquisition Design) and Systematic Functional Linguistics metalanguage (Shleppegrell, 2013) have emerged to provide teachers a way to embed language instruction in their subject content through a balanced literacy approach (Brechtel, 1992; Moore et al., 2018).

However, an approach that has received widespread praise as an effective means of simultaneously addressing second language instruction and content is the Sheltered Instruction Observation Protocol (SIOP) (Guarino et al., 2011). This study seeks to develop a context-specific rating scale to examine the role that training in the SIOP model plays in the perception of self-efficacy among general education teachers regarding their ability to effectively support the needs of ELs, both with and without disabilities, in the general education classroom.

Research Questions

In an effort to develop a tool to measure the perceived self-efficacy and practical knowledge of general education teachers regarding their ability to effectively address the needs of English learner students with and without disabilities, this study will be guided by the following research questions:

1.How many factors (sources of variance) does the Sheltered Instruction Teacher Self-Efficacy Scale (SITSES) demonstrate?

2. Is the Sheltered Instruction Teacher Self Efficacy Scale (SITSES) a valid and reliable instrument?

3. Does Sheltered Instruction Observation Protocol training increase teacher self-

efficacy scores of elementary general education teachers serving ELs with and without disabilities in an inclusive setting, as measured by the SITSES?

Definition of Terms

Below are definitions of terms that are used in this dissertation.

English Learner (EL)/Culturally, Linguistically Diverse Learner (CLDs)*:* a term used in English-speaking countries such as the United States (US) and Canada to describe a person who is learning the English language in addition to his or her native language or any other languages they may speak. ("English-language learner," 2019).

Disability*:* a physical, mental, cognitive, or developmental condition that impairs, interferes with, or limits a person's ability to engage in certain tasks or actions or participate in typical daily activities and interactions (Brown, 2019a).

Disproportionality: is over-representation of minority students identified with a learning disability or other type of disability under the IDEA. When a minority group's numbers in special education are statistically higher than they should be, they are considered disproportionate. (Logsdon, 2020)

Individuals with Disabilities Education Act (IDEA, also known as PL94-142*):* a U.S. federal law that governs how states and public agencies provide early intervention, special education, and related services to children with disabilities. The act addresses the educational needs of children with disabilities from age three to age 18 or 21 in cases that involve 14 specified categories of disability (U.S. Congress, 2007).

Second Language Acquisition (SLA)*:* the process by which people learn a second language. Second-language acquisition (SLA) also refers to the scientific discipline devoted to studying that process. ("Second Language Acquisition," 2019).

Special Education: classes or instruction for students with special needs (any of various difficulties ((such as a physical, emotional, behavioral, or learning disability or impairment)) that causes an individual to require additional or specialized services or accommodations (such as in education or recreation) (Brown, 2019b).

Chapter 2

Review of Literature

To create a scale to measure teacher's perceived self-efficacy to serve ELs with and without disabilities in an inclusive setting, it is important to review the underpinnings of second language (L2) development and disability as they relate to culturally and linguistically diverse learners. As such, the literature review section of this study will begin with an examination of theory and characteristics of L2 development in ELs. The next section will review theory and characteristics of teaching culturally and linguistically diverse students, as well as looking more deeply at the SIOP model. The literature review chapter will conclude with a presentation of research related to teacher efficacy and measurements used to determine and assess teacher self-efficacy.

Theoretical Constructs

The theoretical construct for this examination of teacher self-efficacy (TSE) beliefs is Bandura's (1977) social cognitive theory and psychological construct of self-efficacy (SE). Bandura defines a self-efficacy belief as "the conviction that one can successfully execute the behavior required to produce certain outcomes" (Bandura, 1997, p.3). He posits four sources of information that aid in the formation of SE beliefs: verbal feedback, vicarious experiences, physiological and emotional arousal because of those experiences, and performance mastery experiences (Bandura, 1995). Bandura describes verbal feedback as more persuasive in nature, that aligns with social influences, suggesting that one possesses skills. Vicarious experiences assist one in assessing their skills in relation to others' performance of those skills and the physiological/emotional state that allows one to partially judge their capabilities, strengths, and vulnerabilities. Of these, mastery experiences, according to, Bandura, constitute the strongest

contributor to self-efficacy beliefs because they allow an individual to connect actual experiences to possible future outcomes (Bandura, 1995). However, before examining the research around TSE related to serving culturally, linguistically diverse learners with and without disabilities, it is necessary to understand the theoretical underpinnings of the development of language and learning differences.

Educators face challenges in differentiating whether limited language proficiency in English comprises the barrier to learning or whether it is masking an actual learning disability (Wagner et. al, 2005). For culturally linguistically diverse learners with disabilities to be appropriately supported, educators must determine whether signs of learning disabilities are present before ELs attain English proficiency. An English learner identified as having a learning disability would likely be struggling academically in their first language, as well (Klingner & Eppolito, 2014). Theories about learning, social, and academic language development can provide a framework to consider the relationship between second language acquisition and struggling learners, providing insight to teachers working with culturally, linguistically diverse learners.

Social/Cultural Foundations of Second Language Acquisition

Mikhail Bakhtin's philosophy of language embraced the Vygotskian psychology of language as a social construct (Bryzzheva, 2002). Vygotsky (1978) maintained that to understand the development of mental processes, we need to understand the indicators or signs that mediate them. Bakhtin argued that a word or physical object becomes a sign when it acquires meaning. He stated, "A sign does not simply exist as a given part of reality – it reflects and refracts another reality" (Voloshinov, 1973, p. 10). The idea that a learner's concept of reality helps define their understanding of language relates to culturally and linguistically diverse students' experience

participation in a Eurocentric form of education, implying that two students can engage in a seemingly similar activity for two entirely different reasons, depending on their cultural identity (Bryzzheva, 2002).

According to Vygotsky (1978), "every function in the child's cultural development appears twice: first, on the social level, and later, on the individual level; first, between people (interpsychological) and then inside the child (intrapsychological)" (p. 57). The first level of cultural development can be realized within what is known as Vygotsky's Zone of Proximal Development, defined as "the distance between the actual developmental level as determined by independent problem solving and the level of potential development as determined through problem-solving under adult guidance, or in collaboration with more capable peers" (p. 86).

The work of Vygotsky and Bakhtin influenced the theories and strategies related to SLA (Bryzzheva, 2002; Karpov 2005). The three dominant research approaches to emerge included Behaviorist, Dialogical, and Cognitive-Computational, the last of which has most influenced the field of SLA in the past three decades (Johnson, 2004). The cognitive tradition is based on the hypothetico-deductive method (Harré & Gillett, 1994; Markee 1994) and the more recent information processing-computational model (Bruner 1996; Harré & Gillett 1994). The hypothetico-deductive method allowed researchers to create a series of hypotheses about the internal or mental processes related to SLA based on observation of the external use of language. This method frames the process as a series of linguistic/conversational adjustments made to promote the comprehension of input, which in turn promotes the acquisition of language (Long, 1983). The information processing approach focuses on the human brain as a "computer" that assimilates, processes, and stores rules governing and interpreting the input of linguistic information (Johnson, 2004).

Currently, second language teacher education programs have focused primarily on pedagogical skills and linguistic experience (Faez & Valeo, 2012). However, in the SLA field, a growing body of research stressing the importance of cultural influences and social interactions has suggested a greater emphasis be placed on the socio-cultural context of learning and teaching ELs (Faez & Valeo, 2012; Ortiz & Robertson, 2018; Vygotsky, 1995).

Culture, Language, and Disability

Historically, the research related to disability has been tied to a physical or neurological impairment that impact a child's ability to access educational learning and, subsequently, to fully participate both academically and socially in school (Gindis, 1999). Consistent with his theory of learning, Vygotsky focused his ideas pertaining to disability on the socio-cultural context in which learning occurs. His work on disability and special education has been a more recent addition to the field of research (*Problems of Defectology,* Vygotsky, 1995). His beliefs on children with special needs is surprisingly relevant to the current supporting arguments for educating students with disabilities in a more inclusive environment (Gindis, 1999). According to Vygotsky,

> One must keep in mind that any child with a disability is first a child and only afterwards an impaired child.... One must not perceive in the child with a disability only the defect, the "grams" of the illness and not notice the "kilograms" of health that children possess. From the psychological and pedagogical points of view, one must treat the child with a disability in the same way as a normal child. (Vygotsky, 1995, p. 44)

An important factor in considering disability is the range of perceptions held by various cultures across the globe about people, more specifically students, with disabilities. While more of a dialogue regarding education and support has emerged at the international level, largely due

to the work of the United Nations Convention on the Rights of Persons with Disabilities (Mittler, 2006), varying degrees of acceptance are to be found in many countries regarding the notion of inclusive education for students with disabilities. The model of segregated education for students with disabilities is more common in countries outside the US. Although over 85% of children with disabilities live in developing countries, roughly 2% of these attend any kind of school, while tens of thousands of children with disabilities pend their entire childhood in poorly managed institutions (Mittler, 2006). This trend has continued, it is estimated that there are one billion persons with disabilities worldwide. Of these 10% live in developing countries and of those nearly 90% do not attend school (Sharma, 2015).

While there has been an influx of culturally, linguistically diverse students with identified disabilities whose parents have emigrated to the US seeking a higher quality of education and special education services, a pattern remains of over-identification of ELs referred for special education evaluation (Cummins, 1981; Sullivan, 2011; Sullivan 2017). In the US, students identified as Limited English Proficient (LEP), typically referred to as ELs (ELs), are more likely than their native English-speaking peers to be identified and referred as students in need of special education (Sullivan, 2011). A growing body of research relates to this disproportionate representation of culturally and linguistically diverse learners in special education programs (Artiles et al.,1997; Clark-Gareca et al., 2020; DeValenzuela, et. al., 2006; Sullivan, 2009; Sullivan, 2017). Rueda and colleagues (2002) indicated that Latino children of foreign-born parents were more likely to be identified as disabled, particularly when tested in English. In addition, assessment methods used to determine eligibility for special education were often inappropriately tied to language proficiency, (Rueda et.al.,2002; Swanson et al., 2020).

The Regional Educational Laboratory Northeast & Islands prepared a report based on a qualitative study that examined the processes and challenges associated with identifying learning disabilities in ELs (Sanchez et al., 2010). Challenges stemmed from practitioners' disagreement related to the skills necessary to properly assess and evaluate students and frustration related to the lack of knowledge related to language development and learning disabilities. This issue is highlighted through the court findings of Diana v Board of Education (1970). This class action suit involved a Spanish-speaking student identified as "mentally disabled" (sic) after she was given an intelligence test in English. The court ruled children cannot be placed in special education based on culturally biased tests or those given in languages other than the child's primary language (Diana v State Board of Education 1970). For educators to distinguish between learning challenges due to learning abilities and those due to differing language and culture is not easy. Several researchers have identified the first step in this process as determining a child's language proficiency in both their home language and English (Espinosa & López 2007; Solano-Flores 2008, Swanson et.al., 2020).

One factor to be considered in this assessment of proficiency is the body of research related to the time required for ELs to acquire language used to orally communicate in social situations, which is typically one to two years, as opposed to the time required for them to become cognitively and academically proficient, which is typically five to eight years (Cummins, 1981, 2005, 2008; Klinger & Eppolito, 2014).

Researchers (e.g., Hardin et al., 2007; Markos, 2011; Sullivan 2011) have studied three primary areas regarding ELs and special education. The first is identification of possible factors that influence practitioners in identifying and referring a disproportionate number of ELs to special education, drawing from national and state data to identify patterns (Sullivan, 2011). The

second area of focus is the impact of culture and language on the special education referral and evaluation process. A team of researchers from North Carolina headed by Hardin (2007) examined the referral and evaluation process of ELs to special education to determine how educators addressed the cultural and language factors throughout that process. Finally, the third area of focus is the examination of educators' attitudes and beliefs about culturally and linguistically diverse students. Markos (2011) explored the role of prior knowledge and attitudes of pre-service teachers about ELs in the misidentification of ELs to special education programs.

Patterns and Predictors

Recognizing the rapid growth of culturally diverse learners that include ELs, Sullivan (2011) examined the representation of students identified as ELs in special education disability and placement categories. As part of her research, Sullivan (2011) also investigated the predictive power of Local Education Agency (LEA) characteristics for identification and placement patterns for students identified as ELs. In her review of previous studies, Sullivan (2011) found that, for Latinos in particular, patterns of representation varied substantially at the local level, leading researchers to suggest that further studies be conducted to examine patterns of representation (Klinger et al., 2006). Due to the rapid growth of ELs in school systems, those systems have struggled to meet the needs of the diverse student population (Nguyen, 2012). Creating frameworks for the study, Sullivan (2011) cited four decades of disproportionate numbers of minority students referred to special education or labeled as mentally retarded.

Sullivan (2011) asked to what extent students identified as ELs were disproportionately represented in special education overall and in each of the high-incidence disability categories at the state-level; to what extent the aforementioned state identification patterns were reflected at the LEA-level; to what extent students identified as ELs were placed in the least restrictive

environment at the state and LEA levels; and to what extent disproportionate representation of ELs at the LEA-level was predicted by LEA characteristics.

In discussion of the literature about disproportionality, Sullivan (2011) cited an issue of equity and access in general and special education for culturally, linguistically diverse students that predated the groundbreaking of special education (P.L. 94-142, 1972) which created the formal system of special education in the US. As early as the 1960s, researchers were voicing concerns over the overrepresentation of culturally, linguistically diverse students in classes for the "disabled" (sic) (Dunn, 1968). Included in Sullivan's review was an examination of how disproportionality has been measured historically, along with theoretical perspectives on disproportionality and its connection to poverty (Sullivan, 2011).

Sullivan (2011) conducted an ex post facto study using archival Arizona state data on district general and special education demographic data from 1999 to 2006. A correlational analysis and multiple linear regressions were used to analyze the scope of disproportionality among identified ELs and the relationships between disproportionality and identified district factors, which included instances of white native English-speaking peers represented in special education programs, disability category, type of placement (i.e. percentage of access to general education setting). The study also looked at potential predictors of disproportionate representation, considering district characteristics (district size, student-teacher ratio, teachers certified to teach English as a Second Language [ESL], diversity of teaching staff, percentages of English learner students, minority students, and students qualifying for free/reduced lunch).

Since the study was conducted as a secondary analysis of state data, the sample and selection of variables relied upon state definitions and statutes (Sullivan, 2011). Namely, students were identified as ELs as defined by state law: "a child who does not speak English or

whose native language is not English, and who is not currently able to perform classwork in English" (Arizona Revised Statute 15-751, 2007). Districts included in the study were those reporting enrollment data for ELs. In 1999, the first year of the study, ELs made up 72% of the state's total student population. That number rose to 87% in 2006, at the conclusion of the study.

To examine the likelihood of ELs being identified and placed into special education categories and programs as compared to their white native English peers, Sullivan (2011) used the relative risk ratio (RRR). The RRR, typically used to measure effect size, was used to represent ELs' risk of identification/placement in a given category of special education compared to native English students. Native English students were used as the "English-proficient" reference group. In the state student population used in the study, most of the identified ELs, were Hispanic/Latino (sic), with more than 91% of that group being Spanish speakers of Mexican origin. A positive risk ratio determined that English learner designation was associated with increased likelihood of identification to special education compared to native English students. A negative risk ratio was decreased likelihood. Risk ratio values were set between 0.08 and 1.20 as an acceptable range of risk in correspondence to ratios used in state and federal disproportionality studies. Correlational analyses and multiple regression were used to examine relationships between disproportionality and predictors. The RRRs identified initially were used to order the dependent variables in the models. Significant levels equal to or less than 0.01 were reported to reduce the risk of Type I error.

The RRR data identified an increasing level of disproportionate numbers of ELs in special education in districts across the state from 0.77 in 1999 to 1.19 in 2006. The increased overrepresentation was also present among disability categories. In 2006, ELs were disproportionately identified in three of the high-incidence categories of Speech impairment =

1.82, Mild Mental Retardation =1.63, and Specific Learning Disability =1.30. However, ELs were underrepresented in the emotional disability category=0.12. Examination of the placement of identified ELs demonstrated that at least 51% spent at least 80% of their day in general education, meaning they were more likely to be served in a resource room setting. In 2006, the most current year studied, ELs in general education classrooms with support indicated a risk ratio of 0.80. ELs outside the general education class for less than 21% of their day were disproportionately represented at a risk ratio of 1.02, and ELs accessing general education between 21% and 60% of their day indicated at risk ratio of 1.49.

To assess how well district characteristics predicted disproportionality in special education, Sullivan (2011) used a multiple regression analysis in each of the placement categories and each of the high-incidence disability categories. To rule out a concern of multicollinearity, variance inflation factors were calculated for each predictor. All the variance inflation factors were of acceptable values or <4, eliminating a concern about multicollinearity. The district characteristics examined in Sullivan's study demonstrated the greatest predictive power for overall special education identification with $F(7,134) = 3.58$, $p<0.01$ and placement with limited removal from the general education environment (less than 21%), $F(7,135) = 2.59$, $p<0.01$. Districts that had higher populations of ELs showed less disproportionality in special education placement categories of Specific Learning Disabilities and Speech Language Impairment. Limitations of this study stemmed from the difficulty in analyzing the validity of identification and services where disproportionality existed. Criteria for determining need for services were inconsistent, as was the validity of identification of students as needing special education. Low incidence disorders such as severe cognitive impairments, visual or hearing impairments are sufficiently diagnosed using a medical model of identification-either the

disorder is present or absent. High incidence disorders, such as learning disabilities are more difficult to diagnose using a medical model, these types of disorders are more often determined by a measure of performance and comparative scores on standardized measures whose norm does not always consider cultural and linguistic differences (Wagner et.al.,2005).

After review of the data, Sullivan concluded that ELs were increasingly overrepresented in special education in the state, a finding consistent with the growing body of research and national data regarding ELs in special education (Cummins, 1981, 2000, 2008; Wagner et.al, 2005). Sullivan suggested that further research be conducted focused on the causal relationships among curriculum, instruction, intervention strategies, attitudes and beliefs of educators and teacher training and preparation.

Sullivan's recommendations included the need for patterns of identification of ELs to special education at multiple levels (i.e., district, state, federal) as well as further examination of the methodology, instruction, and professional development of educators providing services to ELs. One challenge with the examination of disproportionality data is the difference from state to state with identification procedures, while the federal law IDEA (2004) lays out criteria for identification, the interpretation and practice of identification and evaluation for special education eligibility varies from state to state. However, in addition to Sullivan's subsequent research (Sullivan & Bal, 2013; Sullivan, 2017) that examined evaluation practices for specific areas of eligibility in a broader geographic area, ongoing research confirms the continued issue of culturally and linguistically diverse students in to special education (NCES, 2018; Skiba et al., 2016).

Hardin, Roach-Scott, & Peisner-Feinberg Study

Due to the increase in young ELs entering the school system in the US over the past five years, Hardin, Roach-Scott, and Peisner-Feinberg (2007) chose to examine the beliefs and practices of early learning educators (administrators and teachers) in relation to the process of referral, evaluation, and placement into special education of preschool ELs. Hardin et al., (2007) cited the frequency with which culturally, linguistically diverse students failed developmental screenings, which caused many of these children to be placed in special education due to educators' difficulties in distinguishing learning differences from cultural and linguistic differences (DeValenzuela et al., 2006). This is congruent with Vygotsky's premise regarding the importance of social and cultural influences on second language learning (Jang & Mangione, 1994).

Hardin et al., (2007) set out to address three questions: The first centered on how cultural and language differences were being addressed during the special education referral, evaluation, and placement process for preschool ELs; the second asked what accommodations were being made to ensure parent participation during the special education referral, evaluation, and placement process; and the third examined whether classroom teachers and special education professionals had been trained in cultural and linguistic practices relevant to the referral, evaluation, and placement process.

To better assess the effectiveness of special education referral, evaluation, and placement of ELs into special education, Hardin et al., (2007) surveyed early learning program administrators and educators in North Carolina, where the growth rate of ELs was highest in the country. The researchers developed two types of measures, a program administrator survey of 45 questions and teacher survey of 26. The surveys included questions regarding types of

assessments, the process of screening, the language of the screening, and assessment and strategies to gather diagnostic information. The researchers also considered the type of early learning programs, geographic regions, and service area type. North Carolina is made of up of three primary regions – Coastal, Mountains, and the Piedmont (center of the state) – so the study included childcare, head start, and public-school early learning programs in the rural and urban/suburban areas of each region. Of the 140 survey participants included in the analysis, 31 completed administrator survey and 109 completed teacher surveys. The respondent group was comprised of 139 females and 1 male, 62.4% White, 32.1% African American, 2.1% Hispanic/Latino, .7% Asian, and 1.4% Alaska Native. Of those respondents who were teachers, 87% of were of Euro/African descent.

The survey was designed to investigate current practices in the referral, evaluation, and placement process for preschool ELs, with questions drafted for each category of services (referral, evaluation, placement) to garner information about both English learner children and their families. Participants were asked to provide personal demographic information (position, gender, race/ethnicity, number of years in early childhood, pre-service and in-service training). Those respondents serving as program directors were asked to identify the type of program they supervised (i.e., public school, kindergarten, private childcare center) and the service area (i.e., rural, urban, suburban) in which their program resided.

Forms of questions included multiple choice and open-ended short answer. In the referral section of the survey, two questions focused on language proficiency in home language, three questions on English language proficiency, and four questions on the preschool screening process. The survey contained four evaluation process questions, including diagnostic assessment instruments used, language in which they were administered, the assessment

administration process, and strategies used to address the needs of English learner children during the diagnostic assessments. Both surveys asked five questions regarding what types of accommodations were offered to parents of ELs throughout the referral, evaluation, and placement process. Additionally, administrators were asked six questions about the skills and training of their teaching staff.

To be included in the study, program administrators and teachers had to meet three criteria, namely that their program enrolled (a) children 3-4 years of age; (b) children who were English language learners; and (c) children with disabilities. The results of the survey were collected and entered in an Excel spreadsheet and verified against the original protocol to ensure accuracy. Open-ended responses were coded to identify themes and patterns. The researchers reported examining data to identify patterns of similarities and differences, noting that they used SPSS to calculate percentages for multiple-choice questions and that some questions allowed multiple answers, which caused total percentages to exceed 100%.

Results indicated that both administrators and teachers reported that determinations of home language and proficiency in home language were made by meetings with parents (A-75.9% and 82.1%, T-65.9% and 60.9%), written forms completed by parents (A-44.8% and 35.7%, T-56.0% and 44.8%), and home visits (A-34.5% and 28.6%, T-45.1% and 37.9%). Both groups also responded consistently about how English language proficiency was determined, with observations in school comprising the most frequent method (A-72.4% and T-80.2%). Regarding questions about what tests were used to determine language proficiency, respondents reported two assessments as employed most often and second most often: the IDEA Oral Language Proficiency Test (IPLT II) and the Miami-Dade Oral Language Proficiency Test. However, a third test identified by both groups was the DIAL-3, which is not a measure of

language proficiency, but rather serves as a developmental screening tool. The teacher group surveyed named the DIAL-3 as the primary instrument used for measuring language proficiency. When questioned about the use made of the language proficiency information, 50% of administrators indicated that results were used for individual planning for students, while nearly 40% of teachers indicated assessment information was used to determine how well children communicated in the language of their family and to determine referral. In regard to the developmental screening process for English language learners, both administrators and teachers responded by naming the DIAL-3 as the primary measure, used once a year, and indicating that both English and Spanish, respectively, were languages used in the screening. Challenges can arise when conducting a bilingual evaluation due to the lack of expertise of educators in the interpretation of assessment results as well as the potential for children who are not proficient in English to perform poorly on the assessment measures (Fernandez & Inserra, 2013).

Questions related to the diagnostic evaluation process yielded similar responses from both administrators and teachers. Both reported that evaluations were administered over 80% of the time in English, with Spanish (i.e., Spanish-DIAL-3) being used second most in the diagnostic assessment process. Classroom observations topped the list of other strategies used in conjunction with the diagnostic assessments. According to both teachers and administrators, diagnostic evaluations were most often administered in the student's home language (Hardin et al., 2007).

When questioned about how Individualized Education Plan (IEP) goals reflected the culturally diverse backgrounds of ELs, 40% of teachers supplied no answer, while 41.2% of administrators indicated that home language and culture were considered. In response to questions about accommodations to assure parent participation in the screening and evaluation

process, both groups shared that sparse information was gathered related to language and culture (Hardin et al., 2007).

Administrators were asked additional questions about preparation and training for teachers and teacher assistants during all phases of the process. Over half of teachers spoke a second language: Spanish, in every case. Local conferences were the primary venue for ongoing staff development and training (Hardin et al., 2007).

In responding to their first research question, based purely on analysis of percentage of responses, Hardin et al. (2007) surmised that no consistent approach was used to determine language proficiency and that heavy emphasis was placed on observational data. Another concern expressed was the use of assessments for language proficiency that were not designed for that purpose, a lack of understanding as to the appropriate uses of screening and assessment tools, and confusion around how the instruments were to be implemented. The authors recommended that more state and federal monies be spent on translating/adapting assessment materials and that these be tested to establish reliability and validity.

Addressing the second research question regarding parent participation, researchers suggested that while some attempt to increase parent involvement had been documented, a need persists for increased training for interpreters regarding early childhood education, special education, and for the referral, evaluation, and screening process of ELs. The authors also expressed a need for more professional development to prepare administrators and teachers to meet the needs of culturally and linguistically diverse students (Hardin et al., 2007).

The reported limitations of this study included a modest sample size of early childhood professionals, many of whom were administrators and teachers from the same programs. Future research completed with larger participant pools and independent samples would increase

likelihood of and range of transferability. Researchers suggested that the results of this and future studies, could aid policymakers and education professionals in the development of more effective strategies and policies regarding the referral, evaluation, and placement process of ELs in special education.

Markos Study

Whereas Sullivan's research explored a growing pattern of disproportionate identification and placement of English language learners into special education, another study (Markos, 2011) looked more specifically at using a process of guided critical reflection (GCR) to address the prior understandings of pre-service teachers about teaching and learning related to ELs. Markos posed research questions regarding the common sense of pre-service teachers about teaching and learning related to ELs; how the use of GCR might transform pre-service teachers' common sense about ELs; and determining teacher's role in creating GCR.

Markos' (2011) review of research described the shortcomings of the state-mandated Structured English Immersion (SEI) coursework – a mandatory program for pre-service teachers in Arizona. The SEI program is designed to prepare pre-service teachers to address the needs of culturally, linguistically diverse students. In contrast to the literature and theorist's belief regarding the importance of using the ELs home culture and first language (L1) to create a cultural congruence between their home and school life, Arizona's English-only approach to SEI coursework places heavy emphasis on instructional strategies and no emphasis on cultural sensitivity, cultural competence, or critical reflection for teachers regarding their practice (Hyun, 1997).

To identify the components related to language learning which teachers need to understand, Markos (2011) examined five areas: teachers' experience with language diversity, a

positive attitude towards linguistic diversity, knowledge related to ELs, second language acquisition knowledge, and skills for simultaneously promoting content and language learning.

Markos (2011) applied the practitioner inquiry method, combining the aspects of teacher research, self-study, and action research. She used a convenience sample made up of students assigned to a course she taught – one of two courses that all pre-service teachers in Arizona were required to take to receive their SEI endorsement and become eligible for a teacher certificate. To assure that students felt comfortable sharing honest opinions, Markos, both teacher and researcher, made it clear that student grades would not be affected by their participation in the study. Out of 24 students, 22 signed consent forms and participated in the study. Within the participant group, 73% were female, 77% were white, and 85% were monolingual English speakers. To assess the group members' previous experience or common sense about ELs, Markos had students respond to the following prompt:

> Explain your personal educational experience, where you went to school K-12 (state, urban/rural, large/small school, etc.) and when you hear the words *English language learner* what comes to mind? (Markos, 2012, p. 46)

Analyzing student responses, Markos identified three central categories (a) respondents with limited exposure to ELs, (b) respondents with a peripheral exposure to ELs, and (c) respondents with a personal experience to ELs during their K-12 education.

Data was also collected from three types of observations:

- Teacher observation notes
- Digital audio recordings
- Outside observer notes

A coding system was designed for her observation notes that distinguished between her actions and her thoughts. The researcher collected in vivo notes during class and ex post facto notes after class, in which she expanded on thoughts captured in the moment. Digital audio recordings were taken via a lapel microphone to capture conversational exchanges with her students. Once a week, an outside observer came to the class and documented the actions of the teacher and students. Throughout the course, Markos (2011) collected student work and formative assessments to include as data in her practitioner inquiry – a systematic and intentional process used by teachers to examine classroom work.

After organizing data, Markos (2011) analyzed how her pre-service teachers used the GCR process as a tool to acknowledge, examine, and transform their thinking about ELs. It was noted that, at the beginning of the course, students maintained a narrow and limited view of ELs, believing them to be Mexican immigrants unable to speak English and maintaining that the best way to teach such learners was through immersion in English with visual aids. However, after being taught the process of GCR, which involves seeing, learning, and acting, the pre-service teachers transformed their common sense regarding not only ELs but also the necessity of participating in an ongoing reflection process to prepare them to support ELs more effectively as they entered the teaching field. Markos suggested that her results will not likely impact Arizona's mandated curriculum for the SEI endorsement portion of the certification process. However, she suggested that positive implications on the delivery of the SEI coursework could be derived from pairing the presentation of the curriculum with the GCR process, which can help teachers gain a better understanding of culturally and linguistically diverse learning.

Limitations of this study included the inability to examine whether pre-service teachers maintained the reflective practice as they continued in their coursework and subsequently to the

teaching field. Another limitation noted was the possibility of a power differential between students and the teacher which may have caused students to tailor their contributions to what they perceived the teacher wanted to see (Markos, 2011). In a study where the teacher is also the researcher there is the potential for relationships or roles to be skewed, where students become more than strictly participants, but whose views or input can create the potential for them to be seen as co-researchers (Looker, 2012). In summary, Markos (2011) suggested that those educators tasked with preparing pre-service teachers be mindful of the need to pair guidance through a critical reflection process with the mandated curriculum. The results of this study suggest the impact of personal beliefs and perceptions that pre-service educators bring with them into the teaching profession, which has been described in more recent research as implicit bias (Weyant, 2019) and the importance of self-reflection to identify this bias.

Stages of Second Language Acquisition

In the process of acquiring a second language, conversational proficiency differs from academic language proficiency. Conversational proficiency is described as basic interpersonal communicative skills (BICS) and academic language proficiency is described as cognitive academic language proficiency (CALP), respectively (Cummins, 2000). Children go through various stages in acquiring a second language. The time required to fully acquire a new language has been estimated at roughly five to seven years; however, new research suggests that the process may require anywhere from seven to ten years (Klinger & Eppolito, 2014). The consensus of linguists is that there are five stages through which ELs pass on their way to language proficiency: (a) preproduction; (b) early production ("telegraphic and formulaic"); (c) speech emergence ("productive language"); (d) intermediate fluency ("continued language development"); and (e) advanced fluency (Klingner & Eppolito, 2014, pg.16). In the pre-

production stage, often called the "silent period," children become quiet and spend most of their time watching and listening, often using non-verbal communication. The next stage, early production, has been described as "telegraphic and formulaic," characterized by children's combining single words to form short phrases or sentences, which are often poorly formed or grammatically incorrect. In the speech emergence or "productive language" phase, vocabulary is increasing and, while many grammatical errors crop up, the speaker can form more complex sentences and communicate in a number of different settings (i.e., classroom, playground, etc.). At the intermediate fluency stage, students typically possess a vocabulary of nearly 6,000 words and can express more complex thoughts and opinions using more complicated sentence structures. This is the point at which students often demonstrate competency in social language use. However, while they may appear to have developed a strong grasp on language usage and grammar, they may still be unable to fully comprehend the expression of complex subject matter concepts. Finally, at the advanced fluency stage, students can function at a level commensurate with native speakers. Reaching this advanced fluency stage of proficiently developed academic language from which students can compete on level ground with their native counterparts can require anywhere from five to ten years (Klingner & Eppolito, 2014).

Instructional Approaches for Second Language Acquisition

With the increased focus on high stakes testing driven by accountability legislation for schools, English language learners, like their native speaking peers, are expected to achieve high academic standards in English and, in many states, to pass end-of-course tests or high school exit exams to graduate (Short et al, 2012; Skiba et al., 2016). ELs have consistently performed below native English speakers, which is not surprising given that most ELs are required to take subject matter area tests in English before they have achieved proficiency in their new language (Short et

al., 2012). In response to this issue, researchers have begun to explore more effective instructional approaches that focus on integrating language development with techniques to make curricular topics more comprehensible (Short et.al., 2012).

Guided Language Acquisition Design

The Guided Language Acquisition Design (GLAD) project was funded by the U.S. Department of Education in the early 1990s. The project developed a model for professional development in language acquisition and literacy (projectglad.com), the focus of which was to promote simultaneously English language acquisition, academic achievement, and cross-cultural skills (Brechtel, 1992).

The GLAD model (Brechtel, 1992) is structured as an integrated, balanced literacy approach incorporating listening, speaking, reading, and writing into all content areas and drawing connections between science, social studies, and literature. The approach posits that language instruction is most successful when it is tied to the meaning of a message or content. Limited research has been conducted on the effectiveness of the GLAD model, which has not been widely implemented in districts across the country.

Systemic Functional Linguistics metalanguage.

Another instructional approach designed to support ELs L2 development and content mastery is the Systemic Functional Linguistics (SFL) metalanguage approach, based on the concept that learners need opportunities for interaction in meaningful contexts coupled with specific language instruction. In a study by Schleppegrell (2013), teachers were trained in implementing SFL using metalanguage which supported situated and contextual language learning, called by for current research in education and L2 acquisition, while also supporting disciplinary goals and activities in English language arts. The construct of SFL draws on the

ways human experience is related through language via three central processes which Schleppegrell (2013) describes as "the *material,* goings-on in the world outside of ourselves; the *mental,* the inner experience of consciousness; and the *relational,* the generalization of experience that identifies and classifies" (p162). The functions and meanings presented in these different process types are realized in different grammatical patterns (Halliday & Matthiessen, 2004). The SFL metalanguage is meaning- and content-based metalanguage providing a vocabulary for raising language awareness that can be linked to the purposes for which language is being used and the goals of the speaker/writer (Schleppegrell, 2013). In examining the use of the SFL approach, researchers concluded that embedding the metalanguage in a particular discipline with a focus on the content to be learned flips the idea of content-based language teaching to one of language-based content teaching, offering new ways to support content learning throughout a school career and across school subjects (Schleppegrell, 2013). One challenge to the widespread use of SFL by teachers working with ELs is that it requires intensive study of complex systems of language and necessary adaptations for pedagogical purposes which is not easily for teachers with limited time to access (Moore et al., 2018).

Sheltered Instruction Observation Protocol

An approach receiving widespread attention is the Sheltered Instruction Observation Protocol (SIOP), developed by a team of researchers in a study funded by the U.S. Department of Education (Guarino et al., 2011). This approach was designed to enable educators to teach content curriculum to students learning a new language using techniques that make content material accessible while also helping students to develop second language skills. SIOP is comprised of 30 features of instruction divided into eight components: Lesson Preparation, Building Background, Comprehensible Input, Strategies, Interaction, Practice and Application,

Lesson Delivery, and Review and Assessment (Echevarria et al., 2011). The model incorporates features recommended for high quality instruction for all students and adds specific features for ELs, such as inclusion of language objectives in each lesson, oral language practice, and the development of background knowledge and academic vocabulary (Guarino et al., 2011).

Developed by a team of researchers over the course of a seven-year study for the Center for Research on Education, Diversity & Excellence (CREDE) and funded by the U.S. Department of Education, the SIOP model evolved in stages (Short, et al., 2011). Over the first four years, the model was developed in collaboration with middle school teachers who field tested a model of sheltered instruction in three urban districts on both coasts of the US. To assess the level of implementation, an observation tool was created for researchers, known as the Sheltered Instruction Observation Protocol (SIOP). By the fourth year of the study, the model had developed into a lesson planning and delivery approach known as the SIOP model (Short et al., 2012).

The next phase of the study developed a four-point scale for each of the 30 features so that a level could be accessed for any lesson. The level ranged from a four (recommended practice) to zero (no evidence of the practice). The study established the validity and reliability of the protocol using independent observers. Statistical analysis revealed an interrater agreement of .90 which, in combination with additional analyses, indicated the SIOP instrument is a highly reliable and valid measure of sheltered instruction (Guarino et al., 2011). A subsequent study by Short et al., (2012) supported these findings.

Next, the researchers examined the model's impact on student achievement. The third phase of the CREDE study was the SIOP writing study, a quasi-experimental study with 19 teachers in the treatment groups and four control teachers that examined whether a significant

difference could be detected in the achievement data for students of the treatment teachers who had received SIOP training vs. students in sheltered classes whose teachers had not received the training (Short et al., 2012). The treatment group consisted of 241 ELs and while the control group consisted of 77 ELs. While the treatment group students scored lower on pre-tests than did the control group, they performed significantly higher on the post-tests. Several studies have confirmed the effectiveness of SIOP (Echevarria et al., 2011; Echevarria et al., 2017; McIntyre et al., 2010; Short et al., 2011; Short et. al, 2012) in positively impacting the achievement of ELs taught by SIOP-trained teachers who implement the model with fidelity (Echevarria et al., 2011).

SIOP has been promoted in districts nationwide to address the high number of culturally, linguistically diverse students entering general education classrooms. SIOP is often introduced via districts' ELs' departments; however, the target of training is general education teachers (Echevarria et al., 2017).

Teacher Self-Efficacy

No all-purpose self-efficacy scale exists that can measure a variety of efficacy beliefs, as items that do not relate to a specific domain of functioning would have little or no explanatory and predictive value (Bandura, 2006). As such, researchers have been challenged in using previously developed Teacher Self-Efficacy (TSE) scales to examine specific factors that are predictive of TSE beliefs related to serving students with specific educational needs, challenging behaviors, or differing cultural or socio-economic backgrounds (Tschannen-Moran et al., 1998).

A number of studies have examined teacher efficacy beliefs and their impact on outcomes such as student achievement (Ashton & Webb 1986; Moore & Essleman, 1992), student motivation (Midgley et al., 1989), the ability to implement classroom management strategies (Woolfolk & Hoy, 1990; Woolfolk et al., 1990), and the ability to work with struggling students

(Ashton & Webb, 1986; Gibson & Dembo, 1984). A consistent theme appeared across this research: teachers with high self-efficacy were open to innovation and willing to try new things (Ghaith & Yaghi, 1997; Guskey, 1988). Alternatively, teachers reporting low self-efficacy felt they had little or no influence on student outcomes and often blamed poor student performance on outside factors (Ashton & Webb, 1986; Bandura, 1997).

Of the studies examining teacher efficacy beliefs and their impact on student outcomes, few have explored the potential factors that impact teacher self-efficacy beliefs about serving specific populations of students. As this present study seeks to develop a tool based on SIOP foundations that can measure teacher self-efficacy beliefs on serving ELs with and without disabilities and can be used to explore what factors (e.g., training, years of experience, second language fluency, etc.) influence teacher's self-efficacy beliefs (e.g., training, years of experience, second language fluency, etc.), it is necessary to review studies examining factors that may influence teacher self-efficacy beliefs. (Tschannen-Moran et. al., 1998).

One study focused on research of pre-service preparation of teachers and general educators' perceptions of their own ability to teach students with a wide variety of learning needs (Barunsteiner & Mariano-Ladipus, 2014). Several studies called out the problem of general educators feeling unprepared or unable to teach students with disabilities (DeSimone & Parmar, 2003; Smith et al., 2000). One of the factors contributing to this perception was the fact that most teacher training programs divide pre-service teachers by the specific certification they pursue, leaving those in the early childhood, elementary, and secondary general education programs with little preparation on how to teach a diverse population of students. Many of these teachers enter classrooms of varied demographics and feel that they lack the training and experience needed to

effectively serve the linguistic, academic, and learning challenges faced by ELs (Tran, 2011; Piazza et al., 2020).

While many studies have examined the connection between teacher self-efficacy and student outcomes, few have explored the influences or predictors of TSE. Bandura and colleagues (1996) indicated that when looking at teacher self-efficacy related to specific student populations, experiences, or outcomes, researchers must develop scales that are context specific. Unlike previous correlational studies, which did not identify causal inferences, one study using a path analysis technique attempted to identify possible factors that may impact teacher beliefs on their self-efficacy in serving identified populations (Brownwell & Pajares, 1996). This study focused on identifying factors that predict a general education teacher's self-efficacy in serving students with learning and behavioral disabilities in the mainstream setting. Brownwell and Pajares (1996) examined the relationships of self-perceptions related to pre-service and in-service education, collegial interactions, support from building administration, and their success in instructing students in the mainstream classroom. The survey instrument they designed, *Working with Diverse Students: The General Educator's Perspective,* was developed with minimal modifications from several scales from previous research (Bandura, 1995; Billingsley, 1995; Morvant & Gersten, 1991; Rosenholtz, 1989). Brownwell and Pajares (1999) conducted a reliability analysis for individual scale items that included: teacher's reported success (Cronbach alpha .81), teacher's efficacy beliefs (Cronbach alpha .90), quality of pre-service preparation (Cronbach alpha .94), quality of in-service preparation (Cronbach alpha .96), General Support from Building Administrator (Cronbach alpha .95), Special Education support from Building Administrator (Cronbach alpha .91), Collegiality with Special Education (Cronbach alpha .89), and Collegiality with General Education (Cronbach alpha .76). After a pilot of the scale was

completed, path analysis techniques were used to examine direct and indirect effects among variables in a predetermined model and for goodness of fit indicators (Brownwell & Pajares, 1999) which proved satisfactory. The researchers found that general education teachers had a greater sense of self-efficacy in their ability to serve students with learning and behavioral difficulties in the general education setting when they had participated in pre-service course work that addresses curricular and instructional strategies for serving students with diverse learning needs and behavior management techniques (Brownwell & Pajares, 1999).

Empirical Evidence

A limited number of studies have examined the self-efficacy of teachers serving ELs, although interest in the topic is growing (Ortiz & Robertson, 2018). Several studies in relation to ELs with disabilities have explored the connection between teacher self-efficacy and the referral and overrepresentation of ELs into special education (Maki et al., 2018; Sullivan 2011; Sullivan 2017). One study conducted by Paneque and Barbetta (2006) examined the teacher self-efficacy beliefs of special education teachers serving ELs with disabilities. The study surveyed 202 elementary special education teachers in an urban district in the southeast region of the US. Researchers used Bandura's (2006) *Guide for Constructing Self-Efficacy Scales* as a model for the development of the Exceptional Children who are ELs (EXCEL) teacher inventory, which also incorporated teacher competency guidelines from the Florida Department of Education for Teachers of English for Speakers of Other Languages (TESOL). The inventory was made up of three sections. Section I included 20 items related to teacher perceptions of the ability to work with students with disabilities from non-English language backgrounds. Items were rated on a nine-point Likert scale. Section II contained three open-ended questions that queried participants about what had been most helpful to them in working with ELs with disabilities, what

recommendations they had regarding future teacher preparation, and what information they felt would be most useful for teachers in the field working with the target population of students. Finally, section III gathered demographic information on participants, which researchers used to correlate to the first two sections of the inventory. This background information included level of schooling (degree obtained), teacher certification, teacher preparation program type, years of teaching experience, status of ESOL endorsement, type of endorsement preparation, and proficiency in languages other than English.

To establish the reliability and validity of the instrument, the researchers used Cronbach's alpha to measure the internal consistency reliability of the instrument. The coefficient alpha was .9419, which demonstrated a highly satisfactory reliability. Researchers calculated correlations and *t*-tests between total teacher efficacy scores and each of the identified demographic variables. To determine a description of the pattern of responses to total efficacy scores, descriptive statistics including frequency distribution and measures of central tendency were calculated. Finally, a multiple regression was employed to analyze which of the teacher variables were the best predictors of teacher efficacy.

Results of the inventory showed that teacher efficacy scores were high. Most of the individual scores ranged from three to nine. The mean and SD were calculated for each of the 20 items. While no statistically significant difference was found in efficacy scores related to years of teacher preparation, years of teaching experience, or socioeconomic status of the students, a statistically significant difference (p=.002) was found in perceived efficacy among participants who reported proficiency in the language of the students. A multiple regression found that proficiency in the home language of students was the predictor of teacher efficacy, p=.001.

The second section of Panneque and Barbetta's (2006) study posed three open-ended questions about what participants thought helped most when working with the target students for themselves, for preservice teachers, and for in-service teachers. The responses were coded and themed, and two themes emerged – organizational issues and teacher issues. Of the two, teacher issues, which included teacher dispositions, need for field-based experience, and teaching skills, emerged as the primary theme. Participants made suggestions related to ways to better support pre-service and in-service teachers, which included a recommendation for more training both in teacher preparation programs and professional development offerings around cultural diversity, testing and evaluation, language development, and second-language acquisition. Those study participants who were proficient in a second language spoke Spanish – the language of a majority of the target students. Panneque and Barbetta (2006) recommended further research, which could include replication of the study in a demographically different school system, or a study of general education classroom teachers who work with the target students, as the current study presents. This study provides insight into teacher perspective in serving ELs, the combination of a quantitative and qualitative survey gave researchers a picture of what factors have the greatest predictor of teacher self-efficacy serving ELs and offered suggestions regarding pre-service and in-service teacher support. However, the study is limited to a specific classification of teachers -special education- in a one region of the United States, which brings into question the ability to generalize findings to a larger population and classification of teachers in a broader geographic area.

Conclusion

The growing attention paid to educational accountability for all students underscores a particularly steep learning curve for culturally and linguistically diverse learners who are new to

the US. Considering the No Child Left Behind (NCLB) act, which has now been reauthorized as, Every Student Succeeds Act (ESSA), teachers feel "challenged" (Dixon et al., 2012, p. 6) to help ELs reach the level of learning complex academic content through English (Dixon et.al., 2012, p. 6). Consistently poor performance by ELs on high stakes testing has spurred a jump to judgment by classroom teachers as to the cause of low-test scores and learning challenges of ELs in content-based classes (Ysseldyke, 2005). The most frequent false assumption is that culturally and linguistically diverse learners have some type of learning disability. Fortunately, a growing field of research around culturally responsive, evidence-based interventions and strategic improvements in practice and policy aims to improve the educational opportunities of ELs in general education by means of specific curricular and instructional language-based strategies that can be embedded into subject matter lessons (Klingner et al., 2005). Additional research is warranted around teacher training on these approaches and the impact that this training may have on teacher self-efficacy. The purpose of the current study is to create an instrument that will examine the perceived self-efficacy ratings of general education teachers who have and have not received training in the Sheltered Instruction Observation Protocol serving ELs with and without disabilities.

Chapter 3

Method

Proposed Research Design

This study employed a correlational research design – a type of quantitative research in which the variables are not manipulated but rather are identified in smaller groups as "factors" that can be statistically related (Vogt, 2005). There are two types of factor analysis, Confirmatory Factor Analysis (CFA) and Exploratory Factor Analysis (EFA). With CFA, there are specific expectations or hypothesis regarding the factors that would be identified, while EFA seeks to identify factors from a set of variables (Vogt, 2005). The exploratory factor analysis was chosen as the purpose of this study was to design a survey instrument.

Purpose

The purpose of this study was to create a reliable and valid survey with which to measure teacher self-efficacy (TSE) in teachers serving a specific student population, namely ELs with and without disabilities.

Participants and Setting

Participants selected were elementary general education teachers from a public school district in the northwestern United States. The district serves over 30,000 students speaking more than 97 languages and, while generally considered suburban, is one of the fastest growing districts in the region, its demographics rapidly shifting due to an influx of students from the international community. The district covers over 72 square miles and is divided into four "learning communities," which include a comprehensive high school and the many elementary and middle schools feeding into that high school. A convenience sample of elementary schools

across the district was selected to administer the survey electronically to general education teachers who serve both students identified as ELs and those who receive special education services in the general education classroom.

Instrumentation

A quantitative study was conducted to accomplish the following: 1) assess the construct validity of the *Sheltered Instruction Teacher Self Efficacy Scale* (SITSES) by collecting data on perceptions, attitudes, and self-concepts of teachers in relation to serving a culturally and linguistically diverse population; and 2) determine whether teachers with and without SIOP training respond differently to the survey. Since the focus of this research study was to create a measure that can be used to examine a possible relationship between general education teachers perceived self-efficacy, pre-service instruction, in-service professional development (SIOP), and experience in using SIOP instructional methodologies to serve ELs with and without disabilities in their classrooms, a survey was developed in alignment with Bandura's *Guide for Constructing Self-Efficacy scales* (2006). According to Bandura (2006), as self-efficacy includes an individual's perceived capabilities, items should include "can do," which measures a judgment of capability, rather than "will do," which measures intentions.

In designing the *Sheltered Instruction Teacher Self Efficacy Scale (*SITSES*)*, the researcher focused on developing items that accurately reflect the construct of the research and target the eight components of the SIOP model likely to exert an impact on the domain of functioning. The study was intended to develop a survey which can be utilized to collect data on attitudes and practice, community needs and or program evaluation and can be helpful in comparing groups (Creswell & Guetterman, 2019). Measures were taken to minimize response

bias; responses were anonymous and confidential, and the scale was identified by a generic title absent the term self-efficacy.

Research questions

1. *How many factors (sources of variance) does the Sheltered Instruction Teacher Self-Efficacy Scale (SITSES) demonstrate?*
2. *Is the Sheltered Instruction Teacher Self Efficacy Scale (SITSES) a valid and reliable instrument?*
3. *Do teachers trained in the Sheltered Instruction Observation Protocol demonstrate higher self-efficacy scores on the SITSES than those who have not been trained?*

Hypothesis.

There will be a statistically significant difference in the scores on the SITSES between teachers with SIOP training and those without SIOP training.

Null hypothesis.

No statistically significant difference will be evident in the scores on the SITSES between teachers with SIOP training and those without SIOP training.

Predictor (independent variable) Trained in SIOP

Criterion (dependent variable) Teacher Self-Efficacy rating (continuous)

Data Analysis

An exploratory factor analysis was used to identify the number of factors and to verify the homogeneity of the scale items. Internal consistency was computed using Cronbach's alpha. Descriptive statistics were computed to provide a description of participants and the pattern of

their responses to the total teacher efficacy score using frequency distributions and measures of central tendency. Exploratory factor analysis is anchored by the basic assumption that there are factors that are present that can help explain the interrelationships in the larger observed variables that exist in development of a measurement instrument (Pett et al., 2003, p 4). To test the hypothesis that there would be a statistically significant difference in self efficacy scores on the SITSES between those teachers with SIOP training and those not trained, an independent sample *t*-test was examined, and further exploration using a non-parametric Mann-Whitney-U test was employed in determining whether to reject the null hypothesis.

Reliability and Validity

As the SITSES was created specifically for the purposes of this research study, reliability and validity were necessarily established through a pilot study. A pilot study was necessary to identify ambiguity or redundancy in survey items and was conducted with small sample sizes larger than 10 but less than 100 (Patten & Newhart, 2018). In this case the pilot study was conducted at two randomly selected elementary school sites. To establish validity and reliability, 21 general education teachers were asked to complete the survey and provide input as to the clarity of the questions. Based on input the questions were revised and the survey was given to a second elementary site. It was necessary to establish content validity to analyze the items in the measure to identify if they measure what they were intended to measure (Patten & Newhart, 2018). To establish content validity, first a review of the literature was conducted relating to the development and implementation of the SIOP model, which included the eight components needed to implement with fidelity. The SIOP teacher self-assessment scale served as a resource to develop component-specific item wording. The initial scale was designed to measure the teacher efficacy related to the eight components making up the SIOP framework, which include

lesson preparation, building background, comprehensible input, strategies, interaction, practice and application, lesson delivery, review and assessment (Guarino et.al., 2011). Next the measure was given to a group of elementary teachers to garner feedback. Finally, to establish face validity, which involves expert review of the scale to confirm that it is measuring the concept it is intended to (Patten & Newhart,2018), the instrument was reviewed by a group of special education program specialists with an English learner endorsement, a teacher who served as a SIOP trainer, a school psychologist working in a school with many ELs, and an English learner teacher. This team, selected due to their expertise in serving ELs and with the SIOP framework, offered recommendations to strengthen the validity of the instrument, resulting in revisions and corrections of identified items. Following their input, the SITSES was administered to a group of 20 classroom teachers to evaluate items for intelligibility and suitability. The scale was provided to teachers, who voluntarily completed it and indicated any suggestions to improve the instrument. Again, revisions were made based on this feedback to eliminate any extraneous factors potentially interfering with the survey's capacity to measure the domain of functioning.

When creating a measure as in this study, it is important to establish the reliability estimate known as internal consistency which looks to establish the consistency of scores within the test (Patten & Newhart, 2018). Internal consistency looks at the correlation of items on a scale and at the extent they measure the same thing (Vogt, 2005). The most common method available to determine internal consistency reliability is Cronbach' s alpha which calculates the variance within an item as well as co-variance between items on a scale (Field, 2017). Cronbach's alpha scores, like other reliability coefficients, range from 0 to 1.0. Scores in the higher range, .70 and above suggest that items more reliably measure the same thing (Patten & Newhart, 2018). Cronbach's alpha was used to measure the internal consistency reliability of the

SITSES instrument, identifying how individual items related to one another and to the instrument as a whole. The Cronbach's alpha for the instrument was calculated at .85, indicating a reliability more than satisfactory. The sum of the teacher efficacy scales was used for data analysis.

Scale Revision

The initial development of the scale and subsequent pilot study was conducted during the 2018-19 school year, with an intention to administer the scale, for the purpose of this research study, in the spring of the 2019-20 school year. However, in the winter of 2020 a *novel coronavirus*, later identified as COVID19, spread across the world causing the World Health Organization and the United States Center for Disease Control (CDC) to declare a global pandemic (Cucinotta & Vanelli, 2020). In March of 2020 public school systems across the United States shifted instruction from in person learning in a classroom to remote learning via computer (Blad, 2020). Due to the unique nature of providing instruction remotely and teachers logging countless hours "on screen", the scale was adjusted in two ways. First to assess teacher's perceived self-efficacy, the nine-point Likert response was streamlined to a five-point Likert response ranging from "Nothing" to "A Great Deal" as it has been recommended by researchers that paring down the number of response options would reduce the frustration level of participant respondents and increase response rate and response quality (Sachdev et al., 2004). Secondly to differentiate between the traditional "in person" learning environment that teachers provided instruction and the "remote" environment that, due to a global pandemic, teachers had to shift their instruction to for the previous eighteen months, an item response line was added to each question to address the different instructional settings, "remote" and "in person". To account for the adjustment of the scale a new Cronbach's alpha was calculated which resulted in an increased score of .939, indicating a stronger reliability in the measure.

Conclusion

The study conducted in this dissertation is an attempt to create a valid and reliable instrument to measure teacher self-efficacy beliefs related to instructional preparation and practice in serving ELs with and without disabilities. The intent of the survey design is to embed SIOP competencies and examine factors, such as pre-service preparation, language fluency, in-service training, and knowledge of SIOP, that contribute to a teacher's self-efficacy. The goal in constructing a valid survey instrument is to enable future researchers to conduct studies to determine areas of focus and training to best prepare teachers to effectively serve ELs in inclusive settings. Additionally, the results of the survey sought to determine if SIOP training contributed to a higher level of self-efficacy ratings from teachers had the training versus those who have not been trained.

Chapter 4

Results

This chapter will present analyses conducted on the measure and two individual data sets representing responses to the *Sheltered Instruction Teacher Self-Efficacy Scale (SITSES)*. To address the original intent and questions of the research study, the reporting of results will focus on the responses labeled "in person". The "remote" vs "in person" results will be addressed but will not serve as the primary focus.

The chapter is organized by first reporting the overall descriptive statistics of the scale. Next, a review of an exploratory factor analysis to determine how many sources of variance are present in the SITSES. Then, analyses of the descriptive statistics for each individual data (SIOP N and SIOP Y) set. Following that will be a review of an added dataset based on learning environment (In Person vs Remote). To conclude, a summary of the most prominent findings is provided.

Scale Descriptive Statistics.

The SITSES was sent electronically to 460 elementary school teachers. The response rate was approximately 10% of teachers responding, resulting in an (n)=45. Females (including transgender females) made up 91.1% of the respondents, while the other 8.9% were male (including transgender male). The majority of respondents, 82.2%, identified as Caucasian, the remaining respondents identified as Asian, 6.7%, Hispanic/Latinx, 4.4%, two or more races, 4.4% and Native American 2.2%. The distribution of participants' years of teaching experience was 35.6% having 0-5 years of experience, 22.2% with 11-15 years of experience, 28.9% with 16 years of experience or more and 13.3% with 6-10 years of experience. In examination of the overall teacher self-efficacy ratings in relation to gender, experience, TSE scores were not remarkable in

comparison to the overall mean score of 71.467. However, in examination of the ethnicity category the mean score was 53 for the two respondents that identified as Hispanic/Latinx, both indicating fluency in Spanish, one having been trained in SIOP and the other had not. It was necessary to verify the assumption of normality and suitability for an exploratory factor analysis (EFA) which was even more critical due to a small sample size. A normal distribution would indicate a skewness and kurtosis value of zero (Field, 2017). The overall in person teacher self-efficacy (TSE) scores indicate a negative skewness of -.922 with a kurtosis of .169 (Table 1). Guidelines for conducting EFA suggest a skewness not exceeding two and kurtosis not exceeding seven are reasonable (Fabrigar & Wegener, 2012).

Table 1

SITSES In Person TSE

N	Mean	SD	Skewness	Kurtosis
45	71.467	7.77	-.922	.169

Exploratory Factor Analysis

While there are several techniques that can be used to extract factors, principal component analysis, principal axis factoring, image factoring, maximum likelihood, alpha factoring and canonical (Williams, Onsman & Brown, 2010), principal solutions tend to be the most common. Exploratory factor analysis (EFA) and Principal Component Analysis (PCA) are often used interchangeably although there is a nuanced difference, as EFA's intent is to find covert constructs underlying a set of evident variables, constructs which are supposed to account for patterns of correlations between the variables. However, PCA's purpose is to reduce data while maintaining as much information from the original data set as possible (Norris &

Lecavalier, 2009). As it serves to summarize many variables into fewer components, a PCA was used in this study.

Suitability of data

Considering the limitation of the sample size, inspection of the correlation matrix, (See Appendix B), showed that all variables had at least one correlation coefficient greater than 0.3 (Pett et al., 2003). Two statistical methods were used to determine adequacy of sample for exploratory factor analysis. Kaiser-Meyer-Olkin (KMO) Measures of Sampling Adequacy is used to see if a sample is acceptable to employ factor analysis. According to Kaiser (1974) a KMO index measure between 0.5 to 1 is considered acceptable. A more detailed interpretation of results suggests ".8 or above, meritorious; .70 or above, middling; .60 or above, mediocre; .50 or above, miserable and below .50, unacceptable" (Hair et al., 2019, p.136). Considering the guidelines, the result of .85 was interpreted as meritorious (see Table 2), indicating the sample adequate for factor analysis.

The second method to identify suitability for factor analysis is Bartlett's Test of Sphericity. When Bartlett's test is conducted, it is expected to be significant at $p<.001$ level. The results of Bartlett's test are listed in Table 2, confirming the data appropriate for factor analysis.

Table 2

KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy		.852
Bartlett's Test of Sphericity	Approx. Chi-Square	523.698
	df	120
	Sig.	.000

Extraction Criteria for Number of Factors

Several techniques were used to decide the number of factors to be extracted, eigenvalue, scree plot and percentage of variance. The accepted rule for extraction is Kaiser's (1974) criteria of eigenvalue > 1 to identify prominent factors. For research in social science the number of factors that account for at least 60% of variance is considered adequate (Hair et al., 2019). Interpretation of the scree plot focuses on the point at which a line through the small eigenvalues breaks, sometimes called the elbow, to indicate the number of factors to be retained (Williams et al., 2010).

Rotation Method

Another question to pose in determining the number of factors to analyze is whether a variable would load on more than one factor. The use of rotation amplifies high item loadings and diminishes low item loadings. There are several types of rotation techniques, orthogonal varimax/quartimax or oblique olbimin/promax, however the main objective of employing a rotation method is to provide ease in interpretation of results (Williams et al., 2010). The Orthogonal Varimax rotation was used to analyze variable loads.

Interpretation

The last step in factor analysis is interpretation, which is used to identify which variables have high factor loadings and can be considered as a single factor, while those variables with low or zero loadings are not identified as a factor. In the analysis for this study, four factors met the criteria identified for extraction, see Table 3 and Figure 1. Each of the four factors had an eigenvalue greater than 1 and cumulatively represented 76.67% of the variance.

Table 3

Cumulative Variances and eigenvalues

Component	Eigenvalue	% of Variance	Cumulative %
1	8.64	53.98	53.98
2	1.46	9.15	63.13
3	1.14	7.15	70.28
4	1.02	6.38	76.67

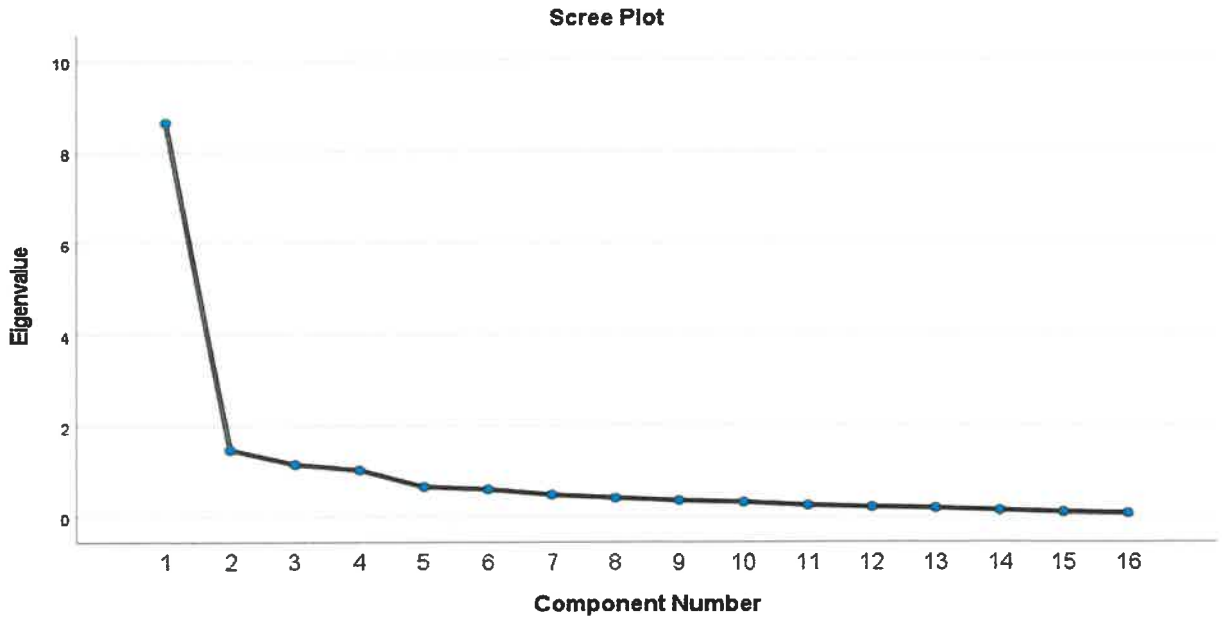

Figure 1. Scree plot

Additionally, the communalities were examined to scan for factors that may not be suitable for analysis. Variables having communalities above .5 are generally acceptable to be retained for analysis (Hair et al., 2019). As such all variables, shown in Table 4, were retained. It is also of note that according to MacCallum and colleagues (1999) "when communalities are all high, sample size will have relatively little impact of quality on solutions, meaning that accurate recovery of population solutions may be obtained using a fairly small sample size." (p. 99). MacCallum and colleagues (1999) also reported that in their review of factor analysis research,

(Harman, 1976; Lawley & Maxwell, 1971; McDonald 198, Mulaik 1972), there were no specific recommendations about sample size.

Table 4

Communalities[a]

Variables	Initial	Extraction	Variable	Initial	Extraction
Q1	1.000	.66	Q9	1.000	.72
Q2	1.000	.87	Q10	1.000	.68
Q3	1.000	.85	Q11	1.000	.82
Q4	1.000	.75	Q12	1.000	.82
Q5	1.000	.72	Q13	1.000	.87
Q6	1.000	.77	Q14	1.000	.76
Q7	1.000	.80	Q15	1.000	.82
Q8	1.000	.73	Q16	1.000	.64

[a]Extraction Method: Principal Component Analysis.

Another area to consider was significant factor loading. It is suggested that an acceptable minimal factor loading is .3 or .4 (Hair et al., 2019). For this study .4 was selected, with the expectation that at minimum each variable would load highly on at least one factor. The design of scale questions was drawn from the eight elements of SIOP (e.g., lesson preparation, building background, comprehensible input, strategies, interaction, practice and application, lesson delivery and review and assessment) and the four component factors (Table 5) that emerged appeared to align with a combination of SIOP elements; 1) Lesson planning and evaluation (SIOP components-*lesson preparation, review and assessment*); 2) Curriculum and instructional strategies (SIOP components-*comprehensible input, strategies*); 3) Prior knowledge and lesson

development (SIOP components-*building background, lesson design*); 4) Student engagement and outcomes (SIOP Components-*interaction, practice and application*).

Table 5

Rotated Component Matrix[a]

	Component			
	1	2	3	4
Q15	**.86**	.16	.20	.12
Q14	**.79**	.26	.21	.13
Q16	**.71**	.33	.03	.19
Q1	**.65**	.21	.19	**.41**
Q11	.23	**.81**	.17	.29
Q4	.22	**.75**	.36	.10
Q10	**.51**	**.64**	.09	.01
Q7	.35	**.61**	.48	.27
Q3	-.05	.21	**.82**	.37
Q2	.38	.31	**.79**	.11
Q13	**.62**	.13	**.65**	.20
Q12	**.44**	**.51**	**.58**	.17
Q5	.04	-.08	.30	**.79**
Q6	.36	.37	.10	**.70**
Q9	**.44**	.35	.10	**.63**
Q8	.16	**.54**	.22	**.61**

[a] Rotation converged in 8 iterations.

Subgroup Descriptive and Inferential Statistics

SIOP Trained vs Not SIOP Trained

To address research question three, descriptive statistics were examined, and inferential statistics were computed on the total teacher self-efficacy scores of teachers that did not receive Sheltered Instruction Observation Protocol (SIOP), (n=12), and those who did (n=33). The mean teacher self-efficacy score of SIOP trained teachers was higher at 73.09 than that of teachers not trained in SIOP with a mean teacher self-efficacy score of 67.00 (Table 6). It is notable that the standard deviation for SIOP trained teachers is lower at 6.57 indicating that on average the scores are closer to the expected value of the set and indicate a stronger reliability (Bland & Altman, 1996).

Table 6

SITSES scores for Teachers not SIOP trained vs. Teachers SIOP trained

Variable	SIOP	*n*	Mean	SD
In Person SE score	No	12	67.00	9.32
	Yes	33	73.09	6.57

The SITSES subgroups data set, shown in Table 7, for SIOP-N indicated a slightly negative skewness of -.366 and kurtosis of -.624, however within acceptable limits. The SIOP-Y data set indicated a -1.001 skewness and a kurtosis of -.306. While the skewness is higher, it does not exceed a value of two which is considered acceptable (Gorsuch, 1983).

Table 7

SITSES Subgroups Skewness and Kurtosis

SIOP Trained	Skewness Statistic	Kurtosis Statistic
No	-.366	-.624
Yes	-1.001	.306

However due to the higher skewness value further examination was conducted to identify potential outliers that could be impacting the data. An outlier was identified as participant 31, shown in Figure 2, in the SIOP-Y data set. The participant whose response represented the outlier of a lower self-efficacy score of 55.00, much lower than the mean of 73.09, identified as Hispanic/Latinx and reported fluency in Spanish. Hair et al. (2019) describes outliers as "those observations that have extremely different values on one or combination of variables" (p. 85). The goal of gathering a representative sample is to capture unique perspectives of a population (Hair et al., 2019) to that end, as the majority of participants in the study identified as Caucasian, it was necessary to maintain the observation of a participant whose ethnicity and perspective was different than the majority. The decision was made to retain the outlier and proceed with the exploratory factor analysis.

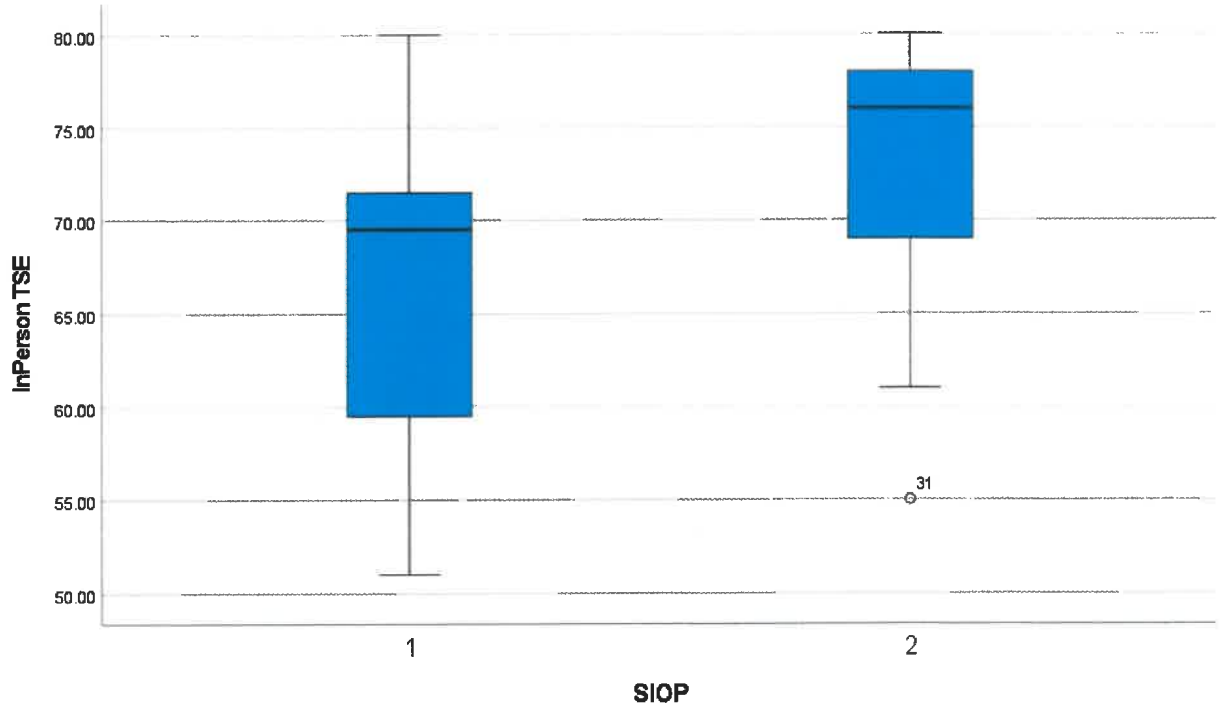

Figure 2. Box Plot of Outliers

There are four assumptions to be met for a parametric test like the independent samples *t*-test, first that the sample distribution is normally distributed, that the data is at least at the interval level, that a homogeneity of variance exists and that scores are independent (Field, 2017). Two of the four assumptions-data measure at the interval level and independent scores-were met, however due to the existence of an outlier in the SIOP-Yes group, the tests of normality were examined to assess the impact of the outlier on a normal distribution. While the scores for those not trained in SIOP were normally distributed as assessed by Shapiro-Wilks, $p = .345$, the self-efficacy scores for those with SIOP training were not normally distributed as assessed by Shapiro-Wilks, $p = .002$ as shown in Table 8.

Table 8
SISTES Subgroup SIOP Not Trained vs SIOP Trained Tests of Normality

		Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	SIOP	Statistic	df	Sig.	Statistic	df	Sig
In Person TSE	No	.209	12	.153	.927	12	.34
	Yes	.186	33	.005	.887	33	.00

[a] Lilliefors Significance Correction

An independent *t*-test was run and Levene's test for equality of variances was examined to establish if the assumption of homogeneity of variances was met. Levene's test is significant at $p \leq .05$ indicating a violation of the assumption of homogeneity of variance exits (Field, 2017), In this case, as shown in Table 10, $p = .166$, indicating that Levene's test is not significant and the assumption of homogeneity of variance is met. According to result of a *t*-test there was a statistically significant difference in teacher self-efficacy scores between teachers with SIOP training and those without SIOP training. SIOP trained teachers demonstrated higher scores (M=73.09, $SE = 1.14$) than teachers without SIOP training (M=67.00, $SE = 2.69$) and $t\ (43) = 2.451$, $p = .018$, demonstrating a medium-sized effect $r = .35$. However, since the normality

assumption was violated for the SIOP-Y group a decision was made to conduct a non-parametric measure to confirm the results of the *t* test. A Mann-Whitney U test was run to determine if there were differences in the in-person teacher self-efficacy (TSE) scores between teachers without SIOP training and teachers with SIOP training. Distributions of the TSE scores for non SIOP trained teachers and SIOP trained teachers were not similar, as assessed by visual inspection. TSE scores for non SIOP trained teachers (mean rank = 16.71) and SIOP trained teachers (mean rank = 25.29) were not statistically significantly different, $U = 273.5$, $z = 1.943$, $p = .052$, using an exact sampling distribution for U (Dineen & Blakesley, 1973).

In Person vs. Remote Instruction

While not the original focus of SITSES, the survey was given after a year of remote instruction due to a global pandemic, so an additional response line was added for teachers to indicate their self-efficacy ratings for remote instruction. Overall, teachers reported higher self-efficacy scores for in person instruction with a mean of 71.467, than the same teachers rating of their self-efficacy teaching remotely with a mean of 55.733 as shown in Table 10. A paired samples *t*-test revealed that there was a statistically significant difference in teacher self-efficacy scores for the in person learning environment compared to the remote environment, M=15.73 (SE=1.65), $t(44) = 9.510$, $p < .001$.

Table 9

SITSES In Person vs Remote TSE

	N	Mean	SD	Skewness	Kurtosis
In Person	45	71.467	7.77	-.922	.169
Remote	45	55.733	9.76	.213	-.340

Remote Instruction SIOP trained vs. non SIOP trained

In examination of the TSE for remote instruction for teachers with SIOP training and those not trained (Table 10) teachers with SIOP training had a slightly higher mean score, however the standard deviation for non SIOP trained teachers was 11.07, suggesting a that there is a wider range of scores from the mean and indicating a lower level of reliability (Bland & Altman, 1996).

Table 10

SITSES scores for Teachers SIOP trained vs teacher not SIOP trained in Remote Instruction

	N	Mean	SD
SIOP Y	33	55.78	9.42
SIOP N	12	55.58	11.07

There was a homogeneity of variances for remote self-efficacy scores between SIOP trained and non-SIOP trained teachers, as assessed by Levene's test for equality of variances (p=.443), however the results of an independent samples t-test indicated that there was not a statistically significant difference in the mean self-efficacy scores, in the remote setting, between teachers with SIOP training(M=55.78, SE=1.64) and those without (M=55.58, SE=3.19), t (43)=.061, p=.951.

Summary of Results

Adjustments to the scale as a result of the pilot study (removal of redundant items), resulted in an adequate Cronbach's alpha of .853. Revisions made subsequent to the pilot study and due to the global pandemic (reduction from a nine-point Likert to a five-point Likert scale, and the addition of a response line for remote instruction) resulted in a stronger internal consistency reliability of .939. Through examination of descriptive statistics and inter-item correlations, the SITSES demonstrated adequate reliability, and suitability for factor analysis.

The exploratory factor analysis showed that four component factors accounted for 76.66% of the variance and that the four factors that emerged aligned with the eight SIOP

framework categories. The factors that were identified were: 1) lesson planning and evaluation (SIOP components-*lesson preparation, review, and assessment*); 2) curriculum and instructional strategies (SIOP components-*comprehensible input, strategies*); 3) prior knowledge and lesson development (SIOP components-*building background, lesson design*); and 4) student engagement and outcomes (SIOP Components-*interaction, practice and application*).

An analysis of the subgroup data set of teachers who were SIOP trained and those who were not, revealed that SIOP trained teachers reported a higher mean self-efficacy score, a mean of 73.09, with a lower standard deviation from the mean at 6.57, than their counterparts that were not trained. A *t*-test indicated that there was evidence of a statistically significant difference between SIOP trained and non-SIOP trained teachers which indicated that the null hypothesis should be rejected, and the alternative hypothesis supported. However as there was not a normal distribution of the SIOP trained teacher group, further investigation was conducted using a non-parametric Mann-Whitney-U test with results suggesting the null hypothesis could not be rejected.

Finally, in examination of the difference in overall teacher self-efficacy scores in an in-person learning environment, with a mean of 71.46, and their overall self-efficacy rating in a remote environment, with a mean of 55.72, a paired samples *t*-test indicated a statistically significant difference. However, an independent samples *t*-test of remote teacher self-efficacy scores for SIOP trained and non SIOP trained revealed no statistically significant difference in their ratings.

Chapter 5

Discussion

The primary focus of this study was to develop an instrument based on the principal framework components of the *Sheltered Instruction Observation Protocol* (SIOP) (Guarino et al., 2011) that sought to measure the perceived self-efficacy of elementary teachers responsible for providing instruction to ELs, with and without disabilities, in an inclusive setting. The study examined the following questions: a) how many factors (sources of variance) would the SITSES demonstrate; b) is the SITSES a valid and reliable instrument; and c) would the teacher self-efficacy scores of teachers with SIOP training be higher than those of teachers without training in SIOP. As the previous chapter summarized the results of data analysis related to these questions, first this chapter will provide an analysis of results of the study, second, will discuss the practical implications of the study, next, the limitations of this study and recommendations for further research are reviewed. Finally, a summary of the most significant findings is presented along with the impact on teacher practice in serving ELs.

Analysis of Results

The first of three research questions asked how many factors would be demonstrated by the scale. In examination of the data the SITSES four components emerged that explained 76.66% of the variance. It was observed that some variables demonstrated high loading on more than one component factor. As the SIOP framework is focused on embedding literacy/language objectives and strategies into content area instruction and there are literacy/language influences across SITSES scale questions, it is not surprising that certain variables loaded highly on more than one factor.

An additional revision of the scale, to address the current environment (i.e. remote instruction) in which teachers were working and to reduce the completion time for the survey by

reducing the rating choices from nine-point Likert to five-point Likert, resulted in a stronger reliability and internal consistency rating, of .94. Finally, addressing the third research question which queried as to the impact of SIOP training on teacher self-efficacy beliefs, the data suggests that those teachers who had received SIOP training demonstrated a higher level of confidence in their ability to design and deliver engaging instruction to students with language learning needs, as evidenced by higher mean self-efficacy scores for in person instructional settings, supported by the results of an independent samples *t*-test. However further examination through a non-parametric Mann-Whitney-U test, indicated that a statistically significant difference between groups was not present and therefore the null hypothesis could not be rejected. Although not a part of the original scale or study, in inclusion of a response line to differentiate learning environments, as a result of the shift to remote instruction due to a global pandemic, examination of teacher self-efficacy ratings for an in person instructional environment versus a remote environment did demonstrate a statistical significance through results of a paired samples *t*-test. This result suggested that teachers, even those with training in SIOP, felt ill-prepared to provide instructional and authentic learning experiences in a virtual classroom setting.

Practical Implications

As discussed in chapter two, there is no one teacher self-efficacy scale that can be utilized to identify predictive factors of teacher self-efficacy as it relates to a specific domain of functioning or to serving students with specific educational needs (Bandura 2006; Tschannen-Moran et al., 2006). There is a growing body of research that has examined teacher self-efficacy beliefs related to serving specific populations of students, their needs, and outcomes (Ashton & Webb, 1986; Brownwell & Pajares, 1999; Moore & Esselman, 1992; Panneque & Barbetta, 2006; Woolfolk & Hoy, 1990). This study was informed by the research of Panneque and

Barbetta (2006) which examined predictors of teacher self-efficacy of special education teachers serving ELs with disabilities. In their study, Panneque and Barbetta (2006) identified a statistically significant difference ($p = .002$) in perceived teacher self-efficacy of teachers who reported proficiency in the language of the students they served. This current study adds to the field of research around measuring teacher self-efficacy related to serving a specific population of students, namely ELs and potential predictors on teacher self-efficacy scores. More specifically that teachers trained in SIOP demonstrate in a higher level of confidence in their ability to effectively serve ELs. As school districts across the nation are faced with a growing population of culturally and linguistically diverse learners (Ortiz & Robertson, 2018), they are challenged to provide teachers with appropriate training to support the needs of students learning English while at the same time presenting these students with core content (Barker & Grassi, 2011). The *Sheltered Instruction Observation Protocol* framework was determined to be an effective tool to provide teachers with a structure and approach to address student language learning and content learning needs at the same time (Guarino et al., 2011). The results from this study showed the SITSES to be a valid and reliable instrument. As the SITSES was developed by incorporating elements from the SIOP components (i.e., lesson preparation, building background, comprehensible input, strategies, interaction, practice and application, lesson delivery, and review and assessment) (Echevarria et al., 2011), it can be utilized by school districts and educational institutions to assess both the need to provide teachers with additional training, and the impact of training that they have received, by examining teacher reported self-efficacy in serving the language and content needs of diverse learners. Additionally, the results of the query into teacher's self-efficacy teaching in a remote setting indicates a need for school districts to

explore more robust training in effective remote instructional practices, particularly in second language development and content learning.

Limitations of the Study

Limitations of this study include sample size and the impact of a significant change in the educational landscape and instructional setting due to a global pandemic. While there is some agreement and general guidelines about sample size and application of factor analysis, there is a range of opinions among researchers on what is acceptable (Gorsuch 1983; MacCallum et al., 1999). In this study, the small sample size was likely impacted by the other limitation of this study which was the stress of a dramatic shift in the educational environment due to a global pandemic, from teachers providing instruction to students in a traditional in person setting to pivoting, with little or no preparation, to a remote style of teaching on a screen for a year, and then within two weeks of the survey being launched, were ordered by the state governor to return to teaching in person, while serving students who remained remote at the same time. Increased time on screens and the need to quickly shift instructional environments was a source of stress for educators and likely impacted their willingness to spend their free time completing an online survey. In addition, due to the requirements of the approved IRB application, surveys sent to teachers needed to be sent from the researcher's university email address which in some cases was identified by the district's server as "spam" and redirected away from teacher's direct email inbox, which impacted the number of surveys received by teachers. Thus, the study proceeded with a small sample size.

Recommendations for Future Research

As the data analysis conducted suggested the SITSES to be a valid and reliable instrument, it would be recommended that the measure be given to teachers during a traditional

(i.e., in person) school year to elicit a larger sample size to confirm the findings of the current study. Additionally, this study focused on teachers serving students in grades K-5, it would also be important to survey teachers serving middle and high school students struggling to grasp English language learning at the same time as attempting to comprehend an increased level of content complexity at the secondary level. It would also be recommended that the further research be done to examine the relationship between SIOP and national focus on developing robust multi-tiered systems of support (MTSS) frameworks (Dulaney, Hallam & Wall, 2013), and the impact of their implementation on referral rates of ELs to special education. As the query to teacher self-efficacy in a remote learning environment was added to the SITSES to account for the sharp change in the instructional environment and since the mean of the teacher self-efficacy scores for remote instruction, 55.73, was notably lower than the mean of the scores in person, 71.67, and resulted in a statistically significant difference through a paired samples *t*-test, another area to explore would be the impact of providing instruction remotely on teachers attitudes and self-efficacy beliefs.

Conclusion

In the fluid state of education with a continuously changing student demographic, it is critical that educators are presented with professional learning, instructional frameworks and strategies to grow their practice in order to authentically and successfully serve the needs of a diverse population of students (Piazza et al., 2020). As school districts and local education agencies seek to identify effective professional growth opportunities to teachers, it is important that they invest in training that is effective in providing teachers with instructional strategies that positively impact student outcomes (Piazza et al., 2020). Having a tool to evaluate teacher's confidence level in their ability to effectively serve their culturally and linguistically diverse

students, is a way to assess the investment of training in the SIOP framework. Additionally, as research suggests teachers that express a higher level of self-efficacy in serving learners with diverse needs are more willing to be innovative in their practice (Ghaith & Yaghi, 1997; Guskey, 1988), it lends support to train them in SIOP. The results of this study suggest that general education teachers who receive training in the *Sheltered Instruction Observation Protocol*, possess a greater level of confidence in serving ELs.

References

Artiles A.J., Trent, S.C., Kuan, L.A. (1997). Learning disabilities research on ethnic minority students: An analysis of 22 years of studies published in selected referred journals. *Learning Disabilities Research & Practice, 12*, 82-91.

Ashton, P.T. & Webb, R.B. (1986). *Making a difference: Teacher's sense of efficacy and Student achievement.* White Plains, NY: Longman.

Arizona Revised Statute (2007) A.R.S.§15-751.

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review, 84*(2),191-215. doi.org/10.1037/0033-295X.84.2.191

Bandura, A. (Ed.). (1995). *Self-efficacy in changing societies.* New York: Cambridge University Press.

Bandura, A. (1997). *Self-efficacy: The exercise of control.* London: Macmillan.

Bandura, A., (2006). Guide for conducting self-efficacy scales. In Pajares, F. & Urdan, T. (Ed.) *Self-efficacy and adolescents* (pp. 307-337) Greenwich, CN: IAP-Information Age Publishing.

Bandura, A., Barbaranelli, C., Caprara, G.V., Pastorelli, C., (1996). Multifaceted Impact of Self-Efficacy Beliefs on Academic Functioning. *Child Development, 67,* 1206 -1222.

Bandura, A., Barbaranelli, C., Caprara, G.V., Pastorelli, C. (2001). Self-efficacy beliefs as shapers of children's aspirations and career trajectories. *Child Development, 72*(1), 187-206.

Barker, H. B., & Grassi, E. (2011). Culturally relevant practices for the special education eligibility process. *AccELLerate, 3*(3), 2-4.

Billingsley, B.S., Pychea, J.N., Smith-Davis, J., Murray, L., & Hendricks, M.B. (1995). *Improving the retention of special education teachers.* (Final report for cooperative agreement H012Q10001). Washington D.C.: Office of Special Education Programs

Blad, E. (2020, March 13). *Millions will be out of school for weeks due to coronavirus. It may not be long enough.* Education Week. https://www.edweek.org/leadership/millions-will-be-out-of-school-for-weeks-due-to-coronavirus-it-may-not-be-long-enough/2020/03

Bland, J. M., & Altman, D. G. (1996). Measurement error. *British Medical Journal, 312*(7047), 1654. doi.org/10.1136/bmj.312.7047.1654

Braunsteiner, M.L. & Mariano-Lapidus, S., (2014). A perspective on inclusion: Challenges for the future. *Global Education Review, 1*(1). 32-43.

Brechtel, M. (1992). *Bringing the whole together: An integrated, whole language approach for the multilingual classroom.* San Diego, CA: Dominie Press.

Bruner, J. (1996). *The culture of education.* Cambridge: Harvard University Press.

Brown, J. (2019a). Disability. In E. M. Sanchez (Ed.), Merriam-Webster. Merriam-Webster. https://www.merriam-webster.com/dictionary/disability

Brown, J. (2019b). Special education. In E. M. Sanchez (Ed.), Merriam-Webster. Merriam-Webster. https://www.merriam-webster.com/dictionary/special education

Brownwell, M.T. & Pajares, F.M. (1999). Teacher efficacy and perceived success in mainstreaming students with learning and behavior problems. *Teacher Education and Special Education, 22*(3), 154 –164.

Bryzzheva, L. (2002). From Vygotsky to Bakhtin: Grand theories and teaching practices. In Diaz Soto, L. (Ed.), *Making a difference in the lives of bilingual/bicultural children* (pp 25-30). New York, NY: Peter Lang.

Clark-Gareca, B., Short, D., Lukes, M., Sharp-Ross, M. (2020). Long-term ELs: Current research, policy, and practice. *TESOL J, 11*(452). doi.org/10.1002/ tesj.452

Creswell, J.W. & Guetterman, T.C., (2019). *Educational research: planning, conducting, and evaluating quantitative and qualitative research* (6th ed.). Upper Saddle River, NJ: Pearson.

Cucinotta, D., & Vanelli, M. (2020). WHO Declares COVID-19 a Pandemic. *Acta bio-medica: Atenei Parmensis*, *91*(1), 157–160. doi.org/10.23750/abm.v91i1.9397

Cummins, J. (1981). The role of primary language development in promoting educational success for language minority students. In California State Department of Education, *Schooling, and language minority students: A theoretical framework* (pp. 3-49). Sacramento, CA: CA Department of Education.

Cummins, J. (2000). *Language, power, and pedagogy: Bilingual children in the crossfire.* Clevedon, UK: Multilingual Matters.

Cummins, J. (2005). A proposal for action: Strategies for recognizing heritage language competence as a learning resource within the mainstream classroom. *The Modern Language Journal, 89*, 585-592. http://www.jstor.org/stable/3588628

Cummins, J. (2008). BICS and CALP: Empirical and theoretical status of the distinction. In B. Street & N. H. Hornberger (Eds.), *Encyclopedia of language and education: 2nd Edition. Volume 2: Literacy* (pp. 71-83). New York: Springer.

DeSimone, J.R., & Parmar, R.S. (2006) Middle school mathematics teachers' beliefs about the inclusion of students with learning disabilities. *Learning Disabilities: Research and Practice, 21*(2), 98-110. doi.org/10.1111/j.1540-5826.2006.00210.x

DeValenzuela, J. S., Copeland, S.R., Qi, C.H., & Park, M. (2006). Examining educational equity: Revisiting the disproportionate representation of minority students in special education. *Exceptional Children, 72*(4), 425-441.

Diana v. California State Board of Education, Civ. No. C-70-37 RFR (N.D. Cal. Jan. 7,1979 and June 18, 1973).

Dineen, L.C., & Blakesley, B.C. (1973). Algorithm AS62: Generator for the sampling distribution of the Mann-Whitney U statistic. *Applied Statistics, 22*, 269-273. doi: 10.2307/2346934

Dixon, L.Q., Zhao, J., Shin, J.Y., Wu, S., Su, J.H., Burgess-Brigham, R., Snow, C. (2012). What we know about second language acquisition: A synthesis from four perspectives. *Review of Educational Research, 82,* 5-60. doi.org/10.3102/0034654311433587

Dulaney, S.K., Hallam, P.R., & Wall, G. (2013). Superintendent's perceptions of multi-tiered systems of support (MTSS): Obstacles and opportunities for school system reform. *AASA Journal of Scholarship and Practice, 10*(2), 30-45.

Dunn, L. M. (1968). Special education for the mildly retarded—Is much of it justifiable? *Exceptional Children, 35*(1), 5–22. https://doi.org/10.1177/001440296803500101

Echevarria, J., Vogt, M.E., & Short, D. (2017). *Making content comprehensible for English language learners: The SIOP model.* New York, NY: Pearson.

Echevarria, J., Richards-Tutor, C., Chinn, V.P., & Ratleff, P.A. (2011). Did they get it? The role of fidelity in teaching ELs. *Journal of Adolescent & Adult Literacy 54*(6), p425-434. doi-org.ezproxy.spu.edu/10.1598/JAAL.54.6.4

Espinosa, L.M, & Lopez, M.L. (2007). *Assessment considerations for young English language learners across different levels of accountability,* A paper prepared for The National Early Childhood Accountability Task Force and First 5 LA. Retrieved November 29, 2016 from https://www.first5la.org/files/AssessmentConsiderationsEnglishLearners.pdf

English-language learner. (2019, May 6). In *Wikipedia.org.* https://en.wikipedia.org/English-language learner

Fabrigar, L.R., & Wegener, D.T. (2012). *Exploratory factor analysis.* New York, NY: Oxford University Press.

Faez, F., & Valeo, A. (2012). TESOL teacher education: Novice teachers' perceptions of their preparedness and efficacy in the classroom. *TESOL Quarterly, 46*(3), 450-471. doi.org/10.1002/tesq.37

Fernandez, N., & Inserra, A., (2013). Disproportionate classification of ESL students in U.S. special education. *TESL-EJ, 17*(2), 1-22.

Field, A., (2017). *Discovering statistics using SPSS.* Thousand Oaks, CA: Sage Publishing

Ford, B.A., Stuart, D.H., & Vokil, S. (2014). Culturally responsive teaching in the 21st century inclusive classroom. *The Journal of the International Association of Special Education 15* (2), 52-62.

Ghaith, G., & Yaghi, H. (1997). Relationships among experience, teacher efficacy, and Attitudes toward implementation of instructional innovation. *Teaching and Teacher Education, 13*(4), 451-458. doi.org/10.1016/S0742-051X(96)00045-5

Gibson, S. & Dembo, M. (1984). Teacher efficacy: A construct validation. *Journal of Educational Psychology, 76,* 569-582. doi.org/10.1037/0022-0663.76.4.569

Gindis, B. (1999). Vygotsky's vision: Reshaping the practice of special education for the 21st century. *Remedial and Special Education, 20*(6), 333-340. doi.org/10.1177/074193259902000606

Gorsuch, R.L., (1983). *Factor analysis* (2nd ed.). Hillsdale, NJ: Erlbaum.

Guarino, A.J., Echevarria, J. Short, D, Schick, J., Forbes, S., & Rueda, R. (2011). The sheltered instruction observation protocol. *Journal of Research in Education, 11*(1), 138-140.

Guskey, T.R. (1988). Teacher efficacy, self-concept, and attitudes toward the implementation of instructional innovation. *Teaching and Teacher Education, 4*, 63-69. doi.org/10.1016/0742-051X(88)90025-X

Hair, J.F., Black, W.C., Babin, B.J., Anderson, R.A., & Tathum, R.L. (2019). *Multivariate data analysis.* Upper Saddle River, NJ: Prentice Hall.

Halliday, M. A. K., & Matthiessen, C. (2004). *An Introduction to Functional Grammar (3rd ed.).* London/Beijing: Edward Arnold/Foreign Language Teaching and Research Press.

Hardin, B., Roach-Scott, M., Peisner-Feinberg, E. (2007). Special education referral, and placement practices for preschool English language learners. *Journal of Research in Childhood Education,* 22(1), 39-54. doi: 10.1080/02568540709594611

Harré, R., & Gillett, G. (1994). *The discursive mind.* Thousand Oaks, CA: Sage.

Harris, B. & Sullivan, A.L. (2017). A framework for bilingual school consultation to facilitate multitier systems of support for English language learners. *Journal of Educational and Psychological Consultation, 27*(3), 367-392. doi.org/10.1080/10474412.2017.1307758

Hyun, E. (1997). Self-examination of one's own ethnicity in the context of teacher preparation for a pluralistic society. Paper presented at the annual conference of the American Educational Research Association, Chicago, IL. (ERIC Document Reproduction Service, No. ED409302)

Jang, Y. & Mangione, P.L. (1994). Transition program practices: Improving linkages between early childhood education and early elementary school. *Southwest Regional Laboratory*, Los Alamitos, CA. (ERIC Document Reproduction Service, No. ED380200)

Johnson, M. (2004). *A philosophy of second language acquisition.* Binghamton, NY: Vail Ballou.

Kaiser, H.F. (1974). An index of factorial simplicity. *Psychometrika, 39,* 31-36. doi.org/10.1007/BF02291575

Kaplan, S. & Leckie, A. (2009). The impact of English-only legislation on teacher professional development: Shifting perspectives in Arizona. *Theory into Practice, 48*(4), 297-303. doi.org/10.1080/00405840903192896

Karpov, Y.V. (2005). *The neo-Vygotskian approach to child development.* New York: Cambridge University Press.

Keller-Allen, C. (2006). *English language learners with disabilities: Identification and other state policies and issues.* Alexandria, VA: Project Forum.

Klingner, J.K., Artiles, A.J., & Bartletta, L.M. (2006). English language learners who struggle with reading: Language acquisition or LD? *Journal of Learning Disabilities, 39*(2), 108-128. doi.org/10.1177/00222194060390020101

Klingner, J.K., Artiles, A.J., Kozleski, E., Harry, B., Zion, S., Tate, W., Durán, G.Z., & Riley, D., (2005). Addressing the disproportionate representation of culturally and linguistically diverse students in special education through culturally responsive educational systems. *Educational Policy Analysis Archives, 13*(38). doi:10.14507/epaa.v13n38.2005

Klingner, J.K., & Eppolito, A.M. (2014). *English language learners: Differentiating between language acquisition and learning disabilities.* Arlington, VA: Council for Exceptional Children.

Logsdon, A. (2020) Disproportionality. In *Disproportionality of Race in Special Education Programs.* https://www.verywellfamily.com/disproportionality-in-special-education-programs-2162684

Long, M. (1985). In Harper, C. deJong, E. (2004). Misconceptions about teaching English language learners. *Journal of Adolescent and Adult Literacy, 48*(2), 152-162.

Looker, S., (2012). Rethinking roles, relationships and voices in studies of undergraduate student writers. *English Teaching: Practice and Critique, 11*(2), 113-127.

Loreman, T., Sharma, U., & Forlin,C., (2013). Do pre-service teachers feel ready to teach in inclusive classrooms? A four-country study of teaching self-efficacy. *Australian Journal of Teacher Education, 38*(1), 27-44. doi:10.14221/ajte.2013v38n1.10

MacCallum, R. C., Widaman, K. F., Zhang, S., & Hong, S. (1999). Sample size in factor analysis. *Psychological Methods, 4*(1), 84–99. doi.org/10.1037/1082-989X.4.1.84

Maki, K.E., Burns, M.K. & Sullivan, A., (2017). Learning disability identification consistency: The impact of methodology and student evaluation data. *School Psychology Quarterly, 32*(2), 254-267. doi.org/10.1037/spq0000165

Markee, N. (1994). Toward an ethnomethodological respecification of second-language acquisition studies. In E. Turene, S. M. Gass, & A. D. Cohen (Eds.), *Research methodology in second-language acquisition.* Hillsdale, NJ: Lawrence Erlbaum.

Markos, A. M. (2011). *Guiding pre-service teachers to critically reflect: Toward a renewed sense about ELs.* (Unpublished doctoral dissertation) Arizona State University, Tempe, AZ.

Markos, A.M. (2012). Mandated to learn, guided to reflect: Pre-service teachers' evolving understanding of English language learners. *Issues in Teacher Education, 21*(1), 39-57.

McIntyre, E., Kyle, D., Chen, C.T., Munoz, M. & Beldon, S., (2010). Teacher learning and ELL reading achievement in sheltered instruction classrooms: Linking professional development to student development. *Literacy Research and Instruction, 49*, 334-351. doi.org/10.1080/19388070903229412

Midgley, C., Feldlaufer, H. & Eccles, J.S. (1989). Change in teacher efficacy and Student self- and task-related beliefs in mathematics during transition to junior high school. *Journal of Educational Psychology, 81*, 247-258. doi.org/10.1037/0022-0663.81.2.247

Mittler, P. (2008). Planning for 2040's: Everybody's business, *British Journal of Special Education, 35*(1),3-10. doi:10.1111/j.1467-8578.2008.00363.x

Moore, J., Schleppegrell, M., & Palincsar, A.S., (2018). Discovering disciplinary linguistic knowledge with ELs and their teachers: Applying systemic functional linguistics concepts through design-based research. *TESOL Quarterly, 52*(4), 1022-1049. doi:10.1002/TESQ.472

Moore, W.P., & Esselman, M.E. (1992, March). *Teacher efficacy, empowerment, and a Focused instructional climate: Does student achievement benefit?* Paper present at the annual meeting of the American Educational Research Association, San Francisco, CA. (ERIC Document Reproduction Service No. ED 350252)

Morvant, M. & Gersten, R. (1995). *Attrition/Retention of urban special education teachers: Multi-faceted research and strategic action planning.* (Final report for cooperative agreement H023Q10001). Washington, D.C.: Office of Special Education Programs.

National Center for Education Statistics. (2010). *Table A-6-1. Number and percentage distribution of 3- to 21-year-olds served under the Individuals with Disabilities Education Act (IDEA), Part B, and number served as a percentage of total public-school enrollment, by type of disability: Selected school years, 1976 –77 through 2007-08.* Washington, DC: Institute of Education Sciences, U.S. Department of Education. Retrieved from http://nces.ed.gov/programs/coe/2010/section1/table-cwd-1.asp

National Center for Education Statistics (2018). *The Condition of Education 2018. Preprimary, elementary and secondary education at a glance xxiv percentage of public-school students who are English language learners.* Washington DC: Institute of Education Sciences, U. S. Department of Education. Retrieved from http://nces.ed.gov/programs/coe/2018

National Center for Education Statistics (2021). *The Condition of Education 2021: English language learners in public schools.* Washington DC: Institute of Education Sciences, U.S. Department of Education. Retrieved from https://nces.ed.gov/programs/coe/indicator/cgf/2021

Norris, M., & Lecavalier, L. (2009). Evaluating the use of exploratory factor analysis in developmental disability psychological research. *Journal of Autism and Developmental Disorders, 40*, 8-20. doi:10.1007/s10803-009-0816-2

Nguyen, H.T., (2012). General education and special education teachers collaborate to support English language learners with learning disabilities. *Issues in Teacher Education, 21*(1), 127-152.

Ortiz, A, A, & Robertson, P.M., (2018). Preparing teachers to serve ELs with language-and/or literacy-related difficulties and disabilities. *Teacher Education and Special Education 2018. 41*(3), 176-187. doi.org/10.1177/0888406418757035

Paneque, O.M., & Barbetta, P.M., (2006). A study of teacher efficacy of special education Teachers of English language learners with disabilities. *Bilingual Research Journal, 30*(1), 171-239. doi.org/10.1080/15235882.2006.10162871

Patten, M.L.& Newhart, M. (2018). *Understanding Research Methods: An overview of the essentials.* New York, Routledge Publishing.

Pett, M.A.,Lackey, N.R., Sullivan, J.J., (2003). *Making sense of factor analysis: Use of factor analysis for instrument development in health care research.* Thousand Oaks, Sage Publishing.

Piazza, S.V., Williams, C., Protacio, M.S., David, V., Tigchelaar, M., & Kuo, H. (2020). Improving instruction for ELs: A professional development study using SIOP. *Journal of Teacher Education and Educators, 9*(3), 383-405.

Rosenholz, S.J. (1989). *Teachers' workplace: The social organization of schools.* New York: Longman.

Rueda, R., Artiles, A.J, Salazar, J., & Higareda, I. (2002). An analysis of special education as a response to the diminished academic achievement of Chicano/Latino students: An update. In R.R. Valencia (Ed.), *Chicano school failure and success: Past, present, and future* (2nd ed., pp. 149 -169). New York, NY: RoutledgeFalmer.

Sachdev, S.B., & Verma, H.V. (2004). Relative importance of service quality dimensions: A multisectoral study. *Journal of Services Research, 4*(1),93-116.

Sanchez, M., Parker, C., Akbayin, B. & McTigue, A. (2010). *Processes and challenges in identifying learning disabilities among students who are English language learners in three New York State districts. (Issues & Answers Report, REL 2010-No. 085)* Washington, DC: Department of Education, Institute of Education Sciences, National Center for Education Evaluations and Regional Assistance, Regional Education Laboratory Northeast and Islands. Retrieved from http://ies.ed.gov/ncee/edlabs.

Schleppegrell, M.J., (2013). The role of metalanguage in supporting academic language development. *Language Learning: A Journal of Research in Language Studies, 63*(1), 153-170. doi:10.1111/J.1467-9922.2012.00742.X

Second-language acquisition. (2019, May 6). In *Wikipedia.org.* https://en.wikipedia.org/wiki/Second-language_acquisition

Sharma, A. (2015). Perspectives on inclusive education with reference to United Nations. *Universal Journal of Educational Research, 3*(5), 317-321. doi: 10.13189/ujer.2015.030502

Short, D.J., Echevarria, J. & Richards-Tutor, C., (2011). Research on academic literacy development in sheltered instruction classrooms. *Language Teacher Research, 15*(3) 363-380. doi.org/10.1177/1362168811401155

Short, D. J., Fidelman, C.G., & Louguit, M. (2012). Developing academic language in English language learners through sheltered instruction. *TESOL Quarterly, 46*(2), 334-361. doi.org/10.1002/tesq.20

Skiba, R.J., Artiles, A.J., Kozleski, E.B., Losen, D.J., & Harry, E.G. (2016). Risks and consequences of oversimplifying educational inequities: A response to Morgan et. al. (2015). *Educational Researcher 45*, 221-225. doi.org/10.3102/0013189X16644606

Smith, S.B., Smith, S.J., & Boone, R. (2000). Increasing access to teacher preparation: The effectiveness of traditional instructional methods in an online learning environment. *Journal of Special Education Technology, 15*(2), 37-46. Retrieved July 16, 2021 from https://www.learntechlib.org/p/91632/.

Solano-Flores, G. (2008). Who is given tests in what language by whom, when and where? The need for probabilistic views of language in the testing of English language learners. *Educational Researcher, 37*, 189-199. doi.org/10.3102/0013189X08319569

Sullivan, A. S. (2009). *Patterns and predictors of English language learner representation in special education.* (Unpublished doctoral dissertation). Arizona State University, Tempe, Arizona.

Sullivan, A. (2011). Disproportionality in special education identification and placement of English language learners. *Exceptional Children, 77*(3), 317-34.

Sullivan, A. & Bal, A. (2013). Disproportionality in special education: Effects of individual and school variables on disability risk. *Exceptional Children, 79*(1), 475-494. doi.org/10.1177/001440291107700304

Sullivan, A. (2017). Wading through quicksand: Making sense of minority disproportionality in identification of emotional disturbance. *Behavioral Disorders, 43*(1), 244-252 doi.org/10.1177/0198742917732360

Swanson, H.L., Kong, J., Petcu, S.D., & Pimentel, M.F.A. (2020). Can difficulties in language acquisition and specific learning disabilities be separated among ELs? *Exceptional Children, 86*(3), 293-309. doi.org/10.1177/0014402919893931

Tran, Y.K. (2011). *Perceptions of teachers' preparedness and efficacy beliefs for teaching English language learners.* (Unpublished doctoral dissertation). University of Texas, Austin, Texas.

Tshannen-Moran, M., Woolfolk Hoy, A., & Hoy, W.K. (1998). Teacher efficacy: Its meaning and measure. *Review of Educational Research, 68*(2), 202-248. doi.org/10.3102/00346543068002202

U.S. Census Bureau. (2008). *U.S. Census Bureau, Population Division.* Retrieved April 11, 2014, from http://www.census.gov/population/

U.S. Census Bureau. (2011). *U.S. Census Bureau, Population Division.* Retrieved April 11, 2014, from http://www.census.gov/population/

U.S. Congress (2007). *Individuals with disabilities improvement act.* IDEA, Pub. L. 108 - 446 Retrieved April 11, 2014 from http://idea.ed.gov/download/statute.html

Vogt, W.P., (2005). *Dictionary of statistics & methodology: A nontechnical guide for the social sciences.* Thousand Oaks, CA: Sage Publishing.

Volosinov, V.N. (1973). *Marxism and the philosophy of language.* Cambridge: Harvard University Press. (Original work published in 1929).

Vygodskaya, G. L. (1999). Vygotsky and problems of special education. *Remedial and Special Education, 20*(6), 330-332. doi.org/10.1177/074193259902000605

Vygotsky, L. S. (1978). *Mind in society.* Cambridge: Harvard University Press. (Original work published in 1934).

Vygotsky, L. S. (1995). *Problemy defectologi* [Problems of defectology]. Moscow: Prosvechenie.

Wagner, R.K., Francis, D.J., & Morris, R.D., (2005) Identifying English language learners with learning disabilities: Key challenges and possible approaches. *Learning Disabilities Research & Practice, 20*(1), 6-15. doi.org/10.1111/j.1540-5826.2005.00115.x

Williams, B., Onsman, A, & Brown, T. (2010). Exploratory Factor Analysis: A five-step guide for novices. *Australian Journal of Paramedicine, 8*(3),1-13. doi:10.33151/ajp.8.3.93

Weyant, J. (2019). Reducing implicit bias toward non-native speakers of English via perspective taking. *Hispanic Journal of Behavioral Sciences, 4*(4), 542-549. doi.org/10.1177/0739986319869321

Woolfolk, A.E., & Hoy, W.K. (1990). Prospective teacher's sense of efficacy and beliefs about control. *Journal of Educational Psychology, 82*(1), 81-91. doi.org/10.1037/0022-0663.82.1.81

Woolfolk, A.E., Rosoff, B. & Hoy, W.K. (1989, April). *Efficacy beliefs and motivational orientations of teachers.* Paper presented at the annual meeting of the American Educational Research Association, San Francisco.

Ysseldyke, J.E. (2005). Assessment and decision making for students with learning disabilities: What if this is as good as it gets? *Learning Disability Quarterly, 28,* 125-128. dxdoi.org/10.2307/1593610.

Zehler, A., Fleishman, H., Hopstock, P., Stephenson, T., Pendrzick, M., & Sapru, S. (2003). *Policy report: Summary of findings related to LEP and SPED-LEP students.* Washington, DC: U.S. Department of Education Office. Retrieved from http://onlineresources.wnylc.net/pb/orcdocs/LARC_Resources/LEPTopics/ED/DescriptiveStudyofServicestoLEPStudentsandLEPStudentsand LEPStudentswithDisabilities.pdf

Appendix A

Sheltered Instruction Teacher Self-Efficacy Scale

1. **How much can you do to create meaningful lesson activities that integrate lesson concepts with language practice opportunities for reading, writing, listening and/or speaking?**

Remote

1	2	3	4	5
Nothing	Very Little	Some Influence	Quite a Bit	A Great Deal

In Person

1	2	3	4	5
Nothing	Very Little	Some Influence	Quite a Bit	A Great Deal

2. **How much can you do to adjust your lessons to the proper level for individual students?**

Remote

1	2	3	4	5
Nothing	Very Little	Some Influence	Quite a Bit	A Great Deal

In Person

1	2	3	4	5
Nothing	Very Little	Some Influence	Quite a Bit	A Great Deal

3. **How much can you do to present concepts explicitly linked to students' background experiences?**

Remote

1	2	3	4	5
Nothing	Very Little	Some Influence	Quite a Bit	A Great Deal

In Person

1	2	3	4	5
Nothing	Very Little	Some Influence	Quite a Bit	A Great Deal

4. **How much can you do to make connections for students between past learning and new concepts?**

Remote

1	2	3	4	5
Nothing	Very Little	Some Influence	Quite a Bit	A Great Deal

In Person

1	2	3	4	5
Nothing	Very Little	Some Influence	Quite a Bit	A Great Deal

5. To what extent can you adjust your speech (enunciation, pace, etc) **to account for students' language proficiency levels?**

Remote

1	2	3	4	5
Nothing	Very Little	Some Influence	Quite a Bit	A Great Deal

In Person

1	2	3	4	5
Nothing	Very Little	Some Influence	Quite a Bit	A Great Deal

6. To what extent can you provide an alternative explanation or example when students are confused?

Remote

1	2	3	4	5
Nothing	Very Little	Some Influence	Quite a Bit	A Great Deal

In Person

1	2	3	4	5
Nothing	Very Little	Some Influence	Quite a Bit	A Great Deal

7. How much can you to do present a variety of tasks that promote higher order thinking skills for all students?

Remote

1	2	3	4	5
Nothing	Very Little	Some Influence	Quite a Bit	A Great Deal

In Person

1	2	3	4	5
Nothing	Very Little	Some Influence	Quite a Bit	A Great Deal

8. To what extent can you craft good questions for your students?

Remote

1	2	3	4	5
Nothing	Very Little	Some Influence	Quite a Bit	A Great Deal

In Person

1	2	3	4	5
Nothing	Very Little	Some Influence	Quite a Bit	A Great Deal

9. To what extent can you get students to learn when English is not their first language?

Remote

1	2	3	4	5
Nothing	Very Little	Some Influence	Quite a Bit	A Great Deal

In Person

1	2	3	4	5
Nothing	Very Little	Some Influence	Quite a Bit	A Great Deal

10. How much can you do to configure groups to support language and content objectives of the lesson?

Remote

1	2	3	4	5
Nothing	Very Little	Some Influence	Quite a Bit	A Great Deal

In Person

1	2	3	4	5
Nothing	Very Little	Some Influence	Quite a Bit	A Great Deal

11. How much can you do to integrate all language skills (i.e. reading, writing, listening and speaking) into activities?

Remote

1	2	3	4	5
Nothing	Very Little	Some Influence	Quite a Bit	A Great Deal

In Person

1	2	3	4	5
Nothing	Very Little	Some Influence	Quite a Bit	A Great Deal

12. To what extent can you provide hands on materials for students to practice new content knowledge?

Remote

1	2	3	4	5
Nothing	Very Little	Some Influence	Quite a Bit	A Great Deal

In Person

1	2	3	4	5
Nothing	Very Little	Some Influence	Quite a Bit	A Great Deal

13. How much can you do to pace your lessons appropriate to all students' ability levels?

Remote

1	2	3	4	5
Nothing	Very Little	Some Influence	Quite a Bit	A Great Deal

In Person

1	2	3	4	5
Nothing	Very Little	Some Influence	Quite a Bit	A Great Deal

14. How much can you do to keep all students engaged 90% to 100% of an instructional period?

Remote

1	2	3	4	5
Nothing	Very Little	Some Influence	Quite a Bit	A Great Deal

In Person

1	2	3	4	5
Nothing	Very Little	Some Influence	Quite a Bit	A Great Deal

15. To what extent can you provide a comprehensive review of key vocabulary and key concepts?

Remote

1	2	3	4	5
Nothing	Very Little	Some Influence	Quite a Bit	A Great Deal

In Person

1	2	3	4	5
Nothing	Very Little	Some Influence	Quite a Bit	A Great Deal

16. How much can you do to assess every student's comprehension and learning of all lesson objectives throughout the lesson?

Remote

1	2	3	4	5
Nothing	Very Little	Some Influence	Quite a Bit	A Great Deal

In Person

1	2	3	4	5
Nothing	Very Little	Some Influence	Quite a Bit	A Great Deal

END OF SURVEY

Please complete Demographic Information on next page

Demographic Information

Gender (you identify as)

- ☐ Female (includes transgender women)
- ☐ Male (includes transgender men)
- ☐ Prefer to self-describe as_____(non-binary, gender fluid, agender, please specify)
- ☐ Prefer not to answer

Ethnicity

- ☐ African American
- ☐ Asian
- ☐ Caucasian, not Hispanic
- ☐ Hispanic/Latino
- ☐ Native American
- ☐ Pacific Islander
- ☐ 2 races or more

Grade(s)/Subjects Taught _____

Years of Experience

- ☐ 0-5
- ☐ 5-10
- ☐ 10-15
- ☐ 15+

Level of Education

- ☐ BA +36
- ☐ MA
- ☐ EDd/Phd

Type of Certification

Elementary Education Secondary Education

Endorsements (check all that apply)

☐Special Ed

☐ EL

☐Other_____

Trained in Sheltered Instruction Observation Protocol (SIOP)

☐Yes

☐ No

Primary Language?_____

Other Language(s) Fluency? _____

Appendix B

Inter-Item Correlation Matrix

	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Q12	Q13	Q14	Q15	Q16
Q1	1.0	.49	.34	.48	.31	.51	.57	.55	.62	.46	.43	.51	.60	.58	.63	.57
Q2	.49	1.0	.67	.54	.26	.45	.71	.49	.47	.47	.48	.82	.79	.53	.47	.43
Q3	.34	.67	1.0	.51	.48	.38	.58	.48	.37	.20	.41	.52	.56	.29	.22	.17
Q4	.48	.54	.51	1.0	.21	.41	.62	.56	.39	.56	.71	.68	.47	.47	.43	.48
Q5	.32	.26	.48	.21	1.0	.49	.25	.39	.41	.14	.29	.32	.33	.27	.20	.16
Q6	.51	.45	.38	.41	.49	1.0	.61	.68	.70	.43	.57	.56	.49	.52	.47	.47
Q7	.57	.71	.58	.62	.25	.61	1.0	.68	.60	.60	.73	.76	.67	.50	.60	.45
Q8	.55	.49	.48	.56	.39	.68	.68	1.0	.59	.34	.59	.58	.43	.38	.37	.45
Q9	.62	.47	.37	.39	.41	.70	.60	.59	1.0	.52	.56	.53	.56	.47	.47	.47
Q10	.46	.47	.20	.56	.13	.43	.60	.34	.52	1.0	.61	.61	.45	.63	.52	.42
Q11	.42	.48	.41	.71	.29	.57	.73	.59	.56	.61	1.0	.63	.47	.48	.41	.53
Q12	.51	.82	.51	.68	.32	.56	.76	.58	.53	.61	.63	1.0	.74	.62	.58	.51
Q13	.60	.79	.56	.47	.33	.49	.67	.43	.56	.45	.47	.74	1.0	.61	.69	.55
Q14	.58	.54	.29	.47	.24	.52	.50	.38	.47	.63	.48	.62	.61	1.0	.75	.62
Q15	.63	.47	.22	.43	.20	.47	.60	.37	.47	.52	.41	.58	.69	.75	1.0	.62
Q16	.57	.43	.17	.48	.16	.47	.45	.45	.47	.42	.53	.51	.55	.62	.62	1.0