2022

CONTROLLING FOR DEVELOPMENTAL LANGUAGE DIFFERENCES IN SPANISH-ENGLISH LEARNERS: A COMPARISON OF THE ORTIZ PICTURE VOCABULARY ACQUISITION TEST AND WOODCOCK MUÑOZ LANGUAGE SURVEY-THIRD EDITION

Amanda Garcia

CONTROLLING FOR DEVELOPMENTAL LANGUAGE DIFFERENCES IN SPANISH-ENGLISH LEARNERS: A COMPARISON OF THE ORTIZ PICTURE VOCABULARY ACQUISITION TEST AND WOODCOCK MUÑOZ LANGUAGE SURVEY-THIRD EDITION

A dissertation submitted in partial fulfillment

of the requirements for the degree of

DOCTOR OF PSYCHOLOGY

to the faculty of the

DEPARTMENT OF PSYCHOLOGY

of

ST. JOHN'S COLLEGE OF LIBERAL ARTS AND SCIENCES

at

ST. JOHN'S UNIVERSITY

New York

by

Amanda García

Date Submitted: September 8th, 2021

Date Approved: February 17th, 2022

Amanda García	Samuel O. Ortiz, Ph.D

ABSTRACT

CONTROLLING FOR DEVELOPMENTAL LANGUAGE DIFFERENCES IN SPANISH-ENGLISH LEARNERS: A COMPARISON OF THE ORTIZ PICTURE VOCABULARY ACQUISITION TEST AND WOODCOCK MUÑOZ LANGUAGE SURVEY-THIRD EDITION

Amanda García

The purpose of this study was to determine whether there is a difference in the vocabulary knowledge of Spanish-speaking English Learners (ELs) when controlling for developmental language norms. This research was essential given the multitude of literature indicating the connection between vocabulary skills and later academic success and the substantial achievement gaps between ELs and English Speakers (ESs). Considering the different normative standards in the Ortiz PVAT and WMLS III English, where the former includes bilingual, exposure-based norms, and the latter does not, this study evaluated the vocabulary knowledge of 27 Spanish-speaking ELs between the ages of 5 and 11 years (M_{age} = 8.22). Participants resided in the Northeastern U.S. with English language exposure ranging from 25% to 91% across their lifespan. This study was conducted remotely due to restrictions on face-to-face interactions resulting from the Coronavirus (COVID-19) global pandemic. Parents served as informants providing demographic and language background via a brief interview. The results proved to be beneficial in expanding research and theory. Specifically, the results supported the hypothesis and demonstrated that using tools that incorporate exposure-based norms (i.e., Ortiz PVAT) offers a more accurate measure of vocabulary knowledge, and those lacking these norms place ELs at great risk of being inappropriately labeled with an educational disability. This was observed through ELs performing within the average range in the Ortiz PVAT and oftentimes low average range in the WMLS III, statistically significant

differences between score means (p <.001), and exceptionally large effect sizes when compared against subtests with non-exposure-based norms (i.e., WMLS III Test 1: Analogies, Test 3: Picture Vocabulary, and Basic English Oral Language Cluster).

DEDICATION

For all those who molded, inspired, and encouraged me, and to all of the children who are English Learners, may your hopes and dreams always be bigger than your fears and doubts. Your success will go beyond anyone's expectations.

ACKNOWLEDGEMENTS

I am thankful to God for giving me strength when I needed it most and blessing me with the village that has facilitated my success thus far. Only He has truly known how much I have wanted to reach this goal and the work and tears I have put into this. I am grateful to my parents for their love and sacrifices and for raising me to have faith in God. Mami, gracias por tu aliento y los sacrificios que has hecho por mí y mis hermanos.

To my sister, Yajaira ("Aya"), and my cousin, Edgar ("Eggie") – thank you for your unconditional love and support. I would not be the woman I am today if it were not for you both playing such a huge role in my life. I have always looked up to each of you for your immense strength and resilience. Aya, you have taught me never to give up. You and mami have shown me what it is to be a strong, independent woman. Eggie, you taught me to strive for greatness at everything I do and never settle for less. To my sister, Jessica, thank you for listening, praying, and being by my side. To my fiancé, Michael, I am truly blessed to have you in my life, and your constant efforts at pushing me through this difficult chapter have made me love you more. Thank you for your ongoing encouragement and your positive attitude. I look forward to a lifetime of accomplishments with you and our future family. To my brother-in-law, Victor, thank you for always lending an ear and supporting me through a number of "bumps" in the road. To my best friend, Monse, thank you for your patience and for always being there for me through many rollercoasters, especially this one. You have always made time to listen to my ideas and helped me in any way necessary. I am blessed to have you as my best friend. To my good friend, Liz, thank you for loving me like a daughter. You believed in me and supported me in every aspect, and I am eternally grateful. Dr. Soye

Zaid-Muhammad (“Dr. Z”) – You have been more than a mentor to me. We have become family. Thank you for empowering me and reminding me that this journey has been about the woman that I am becoming. I love you all. Words cannot express my gratitude. I hope that I have made you all proud.

Thank you, Dr. Samuel O. Ortiz, Dr. Marlene Sotelo-Dynega, and Dr. Dawn P. Flanagan. It has been an honor to have had you all on my committee. Finally, a very special thank you to the children and families who made my study possible. None of this would have become a reality without your willingness and desire to see a change in tools, methods, and strategies used to evaluate English Learners. Thank you.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Introduction

The ethnic, cultural, and linguistic populations within the United States (U.S.) have grown exceedingly for decades (Jones, 2014). In 2017, the U.S. Census Bureau (2018) reported that approximately 49.9% of students in elementary schools were non-Hispanic White. Of notable importance is the decrease in non-Hispanic White students enrolled and the increase of students from other races, particularly Hispanics. It is anticipated that Hispanic Americans will make up approximately 24% of the population by the year 2050 (Ortiz et al., 2008). As the diversity within the United States' population continues to increase rapidly, appropriate assessment methods and tools are necessary in order to make decisions that best support the education of culturally and linguistically diverse students.

As a field, there has been an increase in the importance and use of empirically supported methods alongside theoretically guided conceptualizations and analyses to collect data and guide decision-making processes (Flanagan et al., 2010). The National Association of School Psychologists (NASP) is dedicated to advancing the field of school psychology by advocating and providing services that "promote nondiscriminatory, culturally responsive professional practices and equitable, high-quality education for all of our nation's children and youth" (National Association of School Psychologists [NASP], 2017). It is the responsibility of school psychologists to understand and deliver psychological services that are culturally and linguistically appropriate for students (NASP, 2015; Ortiz et al., 2008).

The purpose of the present study was to determine whether there is a difference in the vocabulary knowledge of Spanish-speaking English Learners (ELs) when controlling

for developmental language norms. In addition, this research aimed to identify fair assessments that could be used to make data-driven decisions concerning the unique educational needs of ELs. Results obtained from this study could enhance assessment measures and increase the field's knowledge of ELs and their language development. This study was necessary given the ample amount of research positing a connection between language abilities and educational success. In addition, a vast amount of literature suggests there are substantial academic achievement gaps between ELs and ESs.

Review of Literature

English Learners (ELs) in the United States

An English Learner (EL) is an individual whose native language is something other than English (Ortiz, 2019) and who is in the process of learning English (Lopez & Velasco, 2017). This population may include not only students who have immigrated to the U.S. but also students born in the U.S. whose home language is something other than English. There is an increase in diversity (Nahari et al., 2017) and ELs in the American public education system (United States Census Bureau [U.S. Census Bureau], 2018; Lopez & Velasco, 2017; Verdugo & Flores, 2007; August et al., 2005). Approximately 4.9 million ELs were enrolled in the U.S. public schools in fall 2016; this was an increase from the 3.8 million EL students enrolled during fall 2010 (National Center for Education Statistics [NCES], 2019). Within the EL student population, there are over 400 languages spoken (Lopez & Velasco, 2017). Spanish is one of the most commonly spoken languages in the U.S. and within the public-school system (NCES, 2019); data indicates that 80% of ELs in the U.S. speak Spanish (Wagner et al., 2005), and Spanish-speaking EL students comprised 76.6% of ELs enrolled in a public school in fall 2016 (NCES, 2019). Additionally, this survey also determined that, as of fall 2016, ELs are more commonly enrolled in urban public schools as opposed to parochial/private schools (NCES, 2019).

Many EL students are academically ill-prepared upon entering high school (Berman et al., 1995). ELs are more likely to perform poorly academically (Skiba et al., 2008) and are at a higher risk of dropping out of school in comparison to other students (Nahari et al., 2017; Rhodes et al., 2005) who are monolingual English Speakers (ESs).

Funding in schools is limited, and education systems have to manage shortages of resources and adequately trained educators of ELs (Verdugo & Flores, 2007). Due to lack of research, inapt educational policies, and the challenges of educators in understanding this population, it remains unclear how to improve methods of educating ELs (Verdugo & Flores, 2007). Despite professional standards and guidelines developed to assess students who are culturally and linguistically diverse (Ortiz, 2019), there is no one set way of educating ELs (Berman et al., 1995).

Historical Litigations Contributing to Improvement in Multicultural Evaluations

Throughout the last 50 years, there have been numerous litigations against the U.S. education system on behalf of families and advocates of culturally and linguistically diverse students. *Arreola v. Santa Ana Board of Education* (1968) and *Diana v. California State Board of Education* (1970) were two court cases that argued the inappropriate placement of Mexican-American children in special education classrooms. Specifically, in *Diana v. California State Board of Education* (1970), Mexican-American students were disproportionately placed in classrooms for students deemed "educationally mentally retarded." As a result of these legal proceedings, assessments of culturally and linguistically diverse students required the use of comprehensive assessments, encouraging involvement from the student's families, and requiring consideration of culture prior to making special education determinations. The rulings resulting from *Diana v. California State Board of Education* (1970) require that EL students are evaluated in their primary language as well as English by a bilingual examiner or bilingual interpreter if one is unavailable. The ruling also required clinicians to reduce their reliance on assessments that are language-oriented (Figueroa, 1989).

Similar to the abovementioned court proceedings, *Guadalupe v. Tempe Elementary School District* (1972) argued against the inappropriate assessment and placement of Hispanic and Native American students in special education classrooms. The outcome of this ruling required students to obtain evaluations that are multidimensional, including an assessment of the student's adaptive skills and a parent interview (Jacob-Timm & Hartshorne 1998, as cited by Esquivel et al., 2007).

The Individuals with Disabilities Education Act (IDEA), formerly known as the Education for All Handicapped Children Act, was enacted in 1975. This federal law mandates that students with disabilities between the ages of 3 and 21 obtain free and appropriate public education (IDEA, 2004; NCES, 2019). IDEA was reauthorized in 2004 in an effort to increase accountability of student progress (IDEA, 2004). Specifically, schools have to utilize evidence-based interventions/methods to support struggling students, coordinate research, and recruit personnel that are appropriate role models for ethnic and linguistically diverse students (IDEA, 2004). More recently, the Every Student Succeeds Act (ESSA), a legislation passed by President Obama in 2015, ultimately led to a shift in oversight over the academic achievement of ELs from the federal government to the state level. Under this federal mandate, schools are required to annually evaluate the English language proficiency of ELs (U.S. Department of Education, n.d.). Some federal regulations within ESSA have been repealed by the U.S. Congress; however, it is unclear how the amendment of this legislation will impact the protective provisions of ELs (Kangas, 2018).

Overrepresentation of ELs as Having Disabilities/Misclassification of ELs

Despite the court rulings and federal mandates, EL students continue to be

overrepresented in special education (Rhodes et al., 2005; Esquivel 1996, as cited by Esquivel et al., 2007; NASP, 2015). In fall 2016, approximately 700,900 or 14.2% of EL students in public schools were identified as having a disability (NCES, 2019). Of the most frequently reported reasons for referring EL students for a special education evaluation, short attention span, poor academic functioning, and oral language factors were most consistent (Rhodes et al., 2005). ELs are often misclassified in special education due to evaluators inappropriately interpreting assessment results (NASP, 2015; Rhodes et al., 2005) or using measures that are inappropriate for the student (Rhodes et al., 2005; Dailor, 2011).

An investigation of educational classifications in the 2017-2018 school year determined that Learning Disability was the most frequently recommended classification, followed by Speech or Language Impairment (NCES, 2019). Due to the nature of a learning disability, the rates at which EL students are classified vary from state to state (Wagner et al., 2005). As per the National Center on Educational Outcomes (NCEO, 2016), approximately 49% of EL students recommended for special education were classified with a learning disability, and 19% were classified with a speech or language impairment. In comparison, approximately 38% of ES students were classified with a learning disability and 17% with a speech and language classification (U.S. Department of Education, n.d.). The enormous number of EL students classified with these disabilities contradicts incidence estimates for severe reading disabilities of approximately 5% in all alphabetic languages (Snowling 2000, as cited by Wagner et al., 2005). Appropriately identifying ELs "with learning disabilities is hampered by a lack of theory and empirical norms that describe the normal course of language and literacy

development for ELs, and the individual, school, and societal factors that relate to that development" (Wagner et al., 2005, p. 13).

Efforts at establishing criteria for a learning disability have been made but were unsuccessful (Wagner et al., 2005). The brain has one phonological processor that keys into the language(s) to which a person is exposed (Wagner et al., 2005); as such, a deficit in the processor can impact language acquisition and reading abilities (Wagner et al., 2005). This also indicates that one cannot experience a true disability in solely one language. If an EL student has established typical phonological skills in his/her native language, then a low score in an English assessment of phonological skills should be explained by insufficient English proficiency and/or lack of instruction and not a disability (Wagner et al., 2005).

Yzquierdo et al. (2004) conducted a study further demonstrating the inappropriate placement of ELs in special education. Specifically, the researchers examined the results of cognitive and language assessments previously administered to EL students in a culturally diverse area during a special education evaluation. This study determined that of the ELs whose cognitive abilities were evaluated in English only, 28% of them were determined eligible for special education services while only 10% were deemed ineligible (Yzquierdo et al., 2004). ELs who obtained evaluations in their native language were not less likely to be deemed ineligible for special education when compared to ELs assessed in English only (Yzquierdo et al., 2004).

Assessment of ELs and the Importance of Utilizing Multiple Assessment Measures

Early research, particularly research conducted prior to the 1990s, was developed on the perception that being culturally and/or linguistically diverse was indicative of

subaverage abilities (Newell & Chavez-Korell, 2017). Many psychological constructs and assessment measures were built upon theories and normed to populations not reflective of (Byrne et al., 2009; Yzquierdo et al., 2004) ELs in the U.S. In order to avoid classifying ELs with disabilities due to their limited proficiency in English, the U.S. Department of Education (2015) obliges school districts to ensure EL students are evaluated in an appropriate language. As indicated by Ortiz (2008), the way cultural and linguistic factors collectively impact assessment performance is poorly researched.

Of great concern when evaluating students, especially ELs, is the use of assessments that are valid and reliable. Dailor (2011) administered a survey to nationally certified school psychologists and determined that one of the most frequently reported ethical concerns for practitioners was assessment-related. Validity may be impacted by language. If a student does not understand task demands because it is presented in a language that he/she does not comprehend, the task is not a valid measure. Although it is impossible to eliminate all biases from an assessment (Ortiz, 2014), further research is needed to ensure that psychological constructs and measures in assessment tools are appropriate for use (Byrne et al., 2009) with students of different cultures whom they are being utilized with.

Assessing culturally and linguistically diverse students appropriately requires the use of multiple assessment methods. “Assessment in only the native language or English can give an incomplete picture of a student’s knowledge, skills, and instructional needs” (Wagner et al., 2005, p. 10). A thorough review of the student’s records, including their cumulative records, recent physical examination, educational history, a home language survey, and any social history evaluations that provide information regarding the family

SocioEconomic Status (SES) and education, is recommended (Elizalde-Utnick & Romero, 2017). Formal and informal observations of the student in his/her natural environments as well as interviews and/or questionnaires with the student and individuals who know the student best (e.g., parents, teachers) are also encouraged (Rhodes et al., 2005; Elizalde-Utnick & Romero, 2017). Work samples and portfolios in both languages should also be reviewed to assist in determining the student's academic skills. Elizalde-Utnick and Romero (2017) indicate that a student's exposure to English instruction and language acquisition should always be considered when assessing culturally and linguistically diverse students. An integrative approach must be used in order to best understand and assess the needs of ELs (Elizalde-Utnick & Romero, 2017) as well as adhering to the legal requisites of former court rulings (*Diana v. California State Board of Education*, 1970; *Guadalupe v. Tempe Elementary School District,* 1972) and federal legislation (i.e., ESSA, 2015).

Language (*Gc*) and Lexical Knowledge (VL)

Comprehension-knowledge or crystallized intelligence (*Gc*) is a broad cognitive ability that is based on the Cattell-Horn-Carroll (CHC) theory (Schrank et al., 2010) and is also known as language-based knowledge (Flanagan & Alfonso, 2017). Language is defined as the use and comprehension "of a spoken (i.e., listening and speaking), written (i.e., reading and writing), and/or other communication symbol system" such as sign-language (American Speech-Language-Hearing Association [ASHA], 2019). Lexical knowledge (VL) is a narrow ability that is defined by the "extent of vocabulary that can be understood in terms of correct word meanings" (Flanagan, 2013; Flanagan & Alfonso, 2017). Language development (LD) is the ability to understand spoken language and is a

general label for language abilities and an intermediate ability between *Gc* and VL (Flanagan & Alfonso, 2017).

Language can be categorized as receptive and expressive (ASHA, 2019). Receptive language is the foundation of general language development. Language includes phonology, vocabulary, semantics, grammar, and pragmatics and involves complex morphology and syntax rules (Collier, 1989; ASHA, 2019). Vocabulary is an important element in improving general reading skills (Luft Baker et al., 2017) and continues expanding throughout an individual's life (Collier, 1989). Critical thinking and language skills can be developed through contextual experiences (Verdugo & Flores, 2007).

Vocabulary development is lifelong (McLaughlin et al., 2000) and cannot be accelerated. The best predictors of reading acquisition and achievement are *Ga* (primarily phonological awareness) and *Gc* (primarily vocabulary; Jongejan et al., 2007). Vocabulary development, particularly receptive vocabulary, significantly influences the development of phonological awareness (Atwill et al., 2010). A student lacking phonological awareness will struggle with understanding an alphabetic writing system (Wagner et al., 2005). Receptive vocabulary refers to an individual's ability to recognize and understand spoken words (Atwill et al., 2010), while expressive vocabulary refers to the production of spoken words (Burger & Chong, 2011).

Cognitive, linguistic, and sociocultural factors influence receptive vocabulary (Atwill et al., 2010). Linguistic and cognitive influences relate to other skills (i.e., memory, retrieval, phonological perception) that directly impact the growth of receptive vocabulary (Atwill et al., 2010), while sociocultural factors reference home, school, and

community influences. Tools utilized in screening receptive [and expressive] vocabulary often use probes to measure the extent and range of students' single-word knowledge (Uchikoshi 2006, as cited by Atwill et al., 2010).

Uchikoshi's (2006) study determined that the literacy experiences of ELs in the home and school play a role in their vocabulary development. Mancilla-Martinez and Lesaux (2011) conducted a longitudinal study annually evaluating the vocabulary development of Spanish-speaking ELs from ages 4.5 to 8 years old and at 11 and 12 years old. Participants received English instruction only in school. Both English and Spanish vocabulary was assessed using the Productive Vocabulary subtest in the Woodcock Language Proficiency Battery-Revised (Mancilla-Martinez & Lesaux, 2011), an expressive vocabulary task requiring students to name pictured objects. Results determined that students whose home language exposure/use (at age 4.5 years) was mostly Spanish and those with equal English and Spanish exposure both had lower English vocabulary scores initially but greater rates of vocabulary growth, in comparison to participants with mostly English language exposure/use (Mancilla-Martinez & Lesaux, 2011). In contrast, these groups initially demonstrated higher Spanish vocabulary scores than participants with mostly English language exposure (Mancilla-Martinez & Lesaux, 2011). Researchers note that despite the increases demonstrated by the Spanish-speaking ELs enrolled in English academic instruction, their gains did not suffice in closing gaps in English vocabulary development.

First and Second Language Acquisition

Language acquisition is the process of learning a language. It occurs subconsciously and is natural learning (Krashen, 1982). It takes 12 years to acquire first

language (Collier, 1989). Second language acquisition is the process of learning another secondary language. Language skills established in one's native language (L1) are transferrable when learning a second language (L2; Verdugo & Flores, 2007). It is recommended that ELs develop appropriate language skills in their native language in an effort to increase their ability to also develop appropriate language in English (August et al., 2002). Acquiring English is a challenging task for ELs and adds to the difficulty of academic learning (Abedi 2011, as cited by Elizalde-Utnick & Romero, 2017). Research indicates that educating ELs in their native language is crucial to learning, understanding, and clarifying content (Verdugo & Flores, 2007).

Second language acquisition involves five stages – preproduction (within the first three months of L2 exposure), early production (within three to six months of L2 exposure), speech emergence (within six months to two years), intermediate fluency (within two to three years), and advanced fluency (Krashen 1982, as cited by Lopez, 2006), which takes approximately five to 10 years. Key characteristics of preproduction include a silent period where the individual has minimal comprehension and expression of L2 and is dependent on modeling and context clues to understand meaning (Krashen, 1982). In early production, the individual utilizes one to three-word phrases (Hearne 2000, as cited by Rhodes et al., 2005). The third stage, speech emergence, involves increased comprehension, expansion of vocabulary, and the use of simple sentences (Hearne 2000, as cited by Rhodes et al., 2005). During the intermediate fluency stage, the individual can have ample, proficient, face-to-face conversations (Hearne 2000, as cited by Rhodes et al., 2005). Lastly, advanced fluency is characterized by well-developed receptive and expressive language skills (Krashen 1982, as cited by Lopez, 2006).

Language Proficiency

Language proficiency is conceptualized in two types – Basic Interpersonal Communication Skills (BICS) and Cognitive Academic Language Proficiency Skills (CALPS; Luft Baker et al., 2017). BICS is the language used daily to converse (Rhodes et al., 2005) and communicate basic needs, while CALPS are language skills pertaining to academic-related information (Verdugo & Flores, 2007; Luft Baker et al., 2017). It takes longer (approximately 4 to 7 years) for someone to establish CALPS than BICS (Verdugo & Flores, 2007).

Language proficiency in the native language (L1) influences the acquisition and development of a second language (L2; Rhodes et al., 2005; Collier, 1989; August et al., 2005). If foundational academic skills were not established in a student's L1 this would affect his/her ability to develop L2 academic skills (Rhodes et al., 2005); Cummins (2016) refers to this interdependence as a Common Underlying Proficiency and explains that this allows for skills, strategies, and concepts to be transferred across languages. Uchikoshi (2006) conducted a study that supports Cummins' theory; Spanish-speaking EL students whose Spanish receptive vocabulary scores were higher at the beginning of the study were likely to have higher English receptive vocabulary scores when compared to those ELs whose Spanish receptive vocabulary scores were lower at the start of the study (Uchikoshi, 2006). When considering L2 acquisition and proficiency, it has been determined that the strongest predictor of academic achievement in L2 is the amount of education obtained in L1 (Thomas & Collier 2002, as cited by Rhodes et al., 2005). Additionally, the student's age when initially exposed to L2 is a crucial variable impacting academic achievement (Collier, 1989).

The language proficiency of ELs should be evaluated in both L1 and L2 (Abedi 2011, as cited by Elizalde-Utnick & Romero, 2017), as it complies with best practices (Ortiz, 2014) and federal mandates. Elizalde-Utnick and Romero (2017) reported that language proficiency assessments of L1 and L2 assist school psychologists with determining the degree of language influence on cognitive and academic measures. Specifically, it is suggested that oral language proficiency and literacy skills in L1 and L2 be evaluated (Elizalde-Utnick & Romero, 2017). Oral language proficiency tasks should assess social language, receptive and expressive abilities, and literacy tasks should evaluate reading and writing skills (Elizalde-Utnick & Romero, 2017). Oral language delays seen in L2 but not in L1 are not inherently related to a disability and, instead, are associated with L2 acquisition (Rhodes et al., 2005).

Language Achievement

ELs have the unique struggle of having to learn English socially while concurrently obtaining academic instruction in this language. It has been assumed that ELs will develop sufficient proficiency in English, resulting in their eventual attainment of academic abilities comparable to ESs (Collier, 1989); however, studies have demonstrated that, when considering standardized tests, even ELs from higher socioeconomic statuses warrant a minimum of five years to earn a 50^{th} normal curve equivalent (Collier, 1989). Educators assume students who establish BICS in L2 will perform well academically (Collier, 1989).

As students enter higher grades, language becomes progressively complex, and there are less available contextual clues (Collier, 1989). In studies conducted with adolescent ELs who have established typical cognitive development in their L1 and who

were instructed in English only, they lagged academically in L2 when compared to their ES peers (Collier, 1989). In order for ELs to successfully acquire L2 and academic achievement, research indicates that EL students must obtain L1 instruction with a gradual introduction of L2 after approximately four to seven years of education (Collier, 1989). Collier (1989) cited various studies demonstrating equivalent, if not better, academic performance by ELs instructed in L1 in comparison to ELs receiving instruction only in L2.

Research conducted by Palermo and Mikulski (2014) found a relationship between the amount of exposure to English and the performance of Spanish-speaking ELs in a standardized receptive vocabulary assessment administered in English. Gámez et al. (2019) conducted a study examining the increase in receptive and expressive language skills of ELs and ESs in kindergarten, resulting from exposure to the academic-based language provided by peers. The results of this study determined a positive relationship between the receptive and expressive language abilities of ELs (Gámez et al., 2019). These studies demonstrate a relationship between the amount and length of English exposure and language development in ELs, which are factors that influence the academic achievement of ELs.

Achievement Gap Between ELs and ESs

In American culture, it is important to read and write in English as it is an essential predictor of educational success (Verdugo & Flores, 2007). There is and always has been a significant gap between the academic achievement of ELs and ESs (Rhodes et al., 2005; August et al., 2005; Snyder & Dillow 2013, as cited by Newell & Chavez-Korell, 2017), especially in reading assessments (Nahari et al., 2017). Early language and

literacy skills contribute to academic success (Hammer et al., 2014). ELs perform one to two standard deviations below ESs on standardized assessments measuring receptive and expressive vocabulary (Hammer et al., 2014).

McLaughlin and colleagues (2000) collected research across varying states and determined that fourth grade ELs have a limited span of vocabulary and lack vocabulary knowledge when compared to same-grade ESs (McLaughlin et al., 2000); additionally, this research determined that the gaps remained consistent over the year (McLaughlin et al., 2000). This research supports the more recently produced data provided by the NCES (2019). Other studies on ELs have demonstrated that reading comprehension is directly impacted by language proficiency in L1 and L2. For example, in a study conducted by Kiefer (2008), Spanish-speaking ELs, who had lower L2 proficiency in kindergarten demonstrated less reading comprehension growth than ELs with higher L2 proficiency.

Data suggests that reading achievement gaps between ELs and ESs increase in higher grades (Nahari et al., 2017; NCES, 2019). As per the NCES (2019), in the 2017-2018 school year, ELs in the fourth grade earned an average reading score that was 37 points below their same-grade ES peers. In this same school year, ELs in the eighth grade earned an average reading score that was 43 points below their same-grade peers (NCES, 2019). Data collected in 2015 determined that twelfth-grade ELs earned an average reading score that was 49 points below their same-grade ES peers (NCES, 2019). This research provides further evidence of the increasing achievement gaps between EL students as they enter higher grade levels.

August et al. (2002) conducted a study where they measured the transfer effects of phonological awareness, word knowledge, word reading, and comprehension from

Spanish to English. At the beginning of the study, the participating students were at the end of the second grade; English reading performance was measured up until the end of fourth grade. Students included monolingual ESs, bilingual (Spanish and English) students receiving English-only instruction, and bilingual (Spanish and English) students who initially received Spanish reading instruction. Results determined that Spanish phonemic awareness, letter identification, and word reading emerged for the students who received Spanish reading instruction (August et al., 2002). Additionally, a positive relationship was found between Spanish passage comprehension measured at the end of second grade and English passage comprehension assessed at the end of fourth grade (August et al., 2002).

Overall, the studies conducted by Palermo and Mikulski (2014), Gámez et al. (2019), and August et al. (2002) suggest that the English language skills of ELs are influenced by the amount of time/exposure spent interacting with ES peers who provide access to diverse vocabulary (Gámez et al., 2019). When developing and norming tools utilized to evaluate ELs, including language proficiency measures, test developers fail to incorporate these variables.

Language Proficiency Assessments Utilized to Assess ELs

Language proficiency assessment tools are typically used to determine language dominance in areas including listening, speaking, writing, reading, and ultimately the student's language of assessment (Elizalde-Utnick & Romero, 2017). When considering the assessment of language proficiency, the goal has been to "determine the degree to which language proficiency in both the native and second languages influences test

performance on cognitive and achievement measures" (Elizalde-Utnick & Romero, 2017, p. 199). There are a few assessment tools utilized to measure language proficiency.

The Woodcock Muñoz Language Survey (WMLS) has been one of the most commonly used language assessment tools. It was originally developed in 1993 and was the first of multiple revisions created to determine CALP levels and listening, speaking, reading, and writing proficiency (Woodcock, Alvarado, Ruef, & Schrank, 2017). At present, this assessment tool is in its third edition. The WMLS III continues to evaluate a student's language proficiency and assists in guiding placement for ELs between the ages of 3 years 0 months and 22 years 11 months old (Woodcock, Alvarado, Ruef, & Schrank, 2017; Houghton Mifflin Harcourt [HMH], 2019). In addition, the WMLS III compares EL students to others of the same grade (HMH, 2019). Two English forms and one Spanish form are available for administration. At this time, there are no published data evaluating the tool.

The Bilingual English-Spanish Assessment (BESA) is a valid and reliable tool that can be utilized to evaluate speech and language abilities (e.g., phonology, syntax, semantics, pragmatics) of children between 4 and 6 years old (Peña et al., 2018). The BESA has three standardized and norm-referenced subtests as well as an activity involving the observation of social language (Peña et al., 2018). It can be administered to ESs or Spanish-speaking ELs (Peña et al., 2018). Although the BESA is reported to be psychometrically sound for use with Spanish-speaking ELs, it is normed only for a limited age group and does not allow for the evaluation of older school-aged children. Further development of tools similar to the BESA, which are standardized to a truly comparable group, are warranted.

Standardization of Assessment Measures and the Concept of True Peers

As reported by Elizalde-Utnick and Romero (2017), ELs are not usually included in the standardization process of assessment instruments. ELs are unique in that they are not comparable to monolingual ESs, but they are also not comparable to monolingual speakers of their native language/country of origin (Rhodes et al., 2005). Once an individual enters the U.S., they are now exposed to English and are technically no longer monolingual speakers. Most assessments are normed to monolingual speakers from other countries (e.g., Mexican, Ecuadorian, Peruvian, Puerto Rican, Dominican, Spaniard monolingual Spanish speakers), who have little or no English exposure (Figueroa, 1989), and distributed in the U.S. for use with ELs. This proves to be problematic as it impacts validity.

Figueroa (1989) explains that the validity of these tools with ELs enrolled in an English educational system is questionable. In other words, an evaluator cannot obtain accurate results with confidence if the population that the assessment utilized is not standardized to the characteristics of the individual participating in the evaluation. Additionally, standardized assessments assume that test results reflect individual differences (Lopez & Bursztyn, 2013); however, "this assumption overlooks the impact of diverse life experiences and relative exposure to the types of tasks required to be performed by the student" (Lopez & Bursztyn, 2013, p. 220). Evaluation results are not indicative of an EL student's ability or weakness but of the EL student's exposure and familiarity with mainstream U.S. culture (Lopez & Bursztyn, 2013).

More than simply comparing students by age, language and cultural knowledge acquisition must be considered when evaluating ELs (Ortiz, 2019). Utilizing assessment

tools with ELs whose culture is not represented within the normed sample causes age-based norms to be incomparable. Lopez and Bursztyn (2013) encourage evaluators to be aware that standardized instruments are not appropriate for ELs, given the lack of validity for the population.

Rhodes et al. (2005) recommend evaluators compare ELs with other ELs in the same grade, while Collier (1989) discusses two important measures to consider when assessing L2 acquisition - "age on arrival" and "length of residence." Age on arrival is indicative of the age when the individual arrived in the country and was initially exposed to L2 (Collier, 1989). Length of residence references the amount of time – months or years – of exposure to L2 (Collier, 1989).

As noted in prior mentioned studies, including Palermo and Mikulski (2014) and Gámez and colleagues (2019), the amount of time an individual is exposed to a language influences their vocabulary development. In addition, the amount of exposure to mainstream U.S. culture influences test performance. Ortiz (2019) argues that there are levels of cultural knowledge that are expected developmentally; this theory is supported by mounds of evidence in research and practice with ELs demonstrating comparable performance to ESs in nonverbal subtests and large differences in performance within verbal tasks. Ortiz (2018a; 2018b) attempts to further equalize the assessment of ELs by comparing them to other ELs, of the same age, who have been exposed to English for the same amount of time; Ortiz (2019) deems this to be a "true peer" comparison group.

Importance of Study to School Psychologists

School psychologists are responsible for researching and delivering the most appropriate services for students and making decisions based on data. Cultural

competence is an essential tool that plays a role in all areas of psychological service delivery (Ortiz et al., 2008). NASP and the American Psychological Association (APA) highly value cultural competence and promote its skill development in training programs and within professional activities (Ortiz et al., 2008; NASP, 2017).

Present Study

This study was concerned with understanding the vocabulary knowledge of ELs. Vocabulary knowledge is one of the best predictors of reading acquisition. More specifically, the goal of this study was to determine whether there is a difference between the vocabulary knowledge of Spanish-speaking ELs when controlling for developmental language differences. Research is needed to better understand ELs and the tools used to guide special education eligibility and recommendations. This research also sought to identify equitable assessment measures for ELs that could be utilized to make data-driven decisions concerning the unique educational needs of ELs and special education referrals. Information obtained from this study could increase assessment knowledge and literature of ELs as well as improve the field's insight of developmental language norms in ELs, which could further allow school psychologists to enhance their skills in working with this population. This research was essential given the multitude of literature indicating the connection between language/vocabulary skills and later academic success and the substantial academic achievement gaps between ELs and ESs.

Research Question

This study intended to answer the following question:

1. Will Spanish-speaking English Learners (ELs) yield different vocabulary scores as measured by the Ortiz PVAT than the WMLS III English?

Hypothesis

The researcher hypothesized:

1. Assuming the WMLS III's norming sample is potentially discriminatory because it is based on monolingual, non-exposure norms, Spanish-speaking ELs would perform significantly differently (deviation below the normative average is likely) on the WMLS III than on the Ortiz PVAT due to the latter's use of bilingual, exposure-based norms.

Method

Participants

Participants were recruited from the Northeastern part of the U.S. Spanish-speaking families were recruited through flyers that were handed out in public areas such as parks. The researcher was available to discuss the study either in-person, at the time of delivering the flyers, or via telephone, depending on the parent(s) availability. These families were asked to also share the study information with other parents of Spanish-speaking ELs with whom they were familiar.

A total of 28 Spanish-speaking ELs between the ages of 5 and 11 years old participated in the study. Participants were identified as Spanish-speaking ELs if they were exposed to Spanish since birth and continued to be exposed to that language in their upbringing. The participants' English language exposure ranged from 25% to 91% across their lifespan. Parents served as informants providing demographic and language background information via a brief interview.

A total of 27 participants – 12 males (44.4%) and 15 females (55.6%) were included in the study. A 5-year-old male who recently (one month before testing) arrived in the U.S. from Guatemala was not included in the study, given that his performance scores in the WMLS III fell below a standard score of 40, which was less than the minimum score allowed. As there were insufficient participants with a similarly brief duration of English language exposure, this participant's data presented as an outlier and was excluded from the final analysis.

Once the outlier was removed, the performance of 27 participants was analyzed. The mean participant age was 8.2, and the median and mode age was 9.0 years old.

Participants ranged in grades from prekindergarten to 5th grade, while one of the male participants (3.7%) was never enrolled in school. Specifically, there was one participant (3.7%) enrolled in prekindergarten, two participants (7.4%) enrolled in kindergarten, six participants (22.2%) enrolled in first grade, and two participants (7.4%) enrolled in second grade. In addition, there were five participants (18.5%) enrolled in third grade, five (18.5%) enrolled in fourth grade, and five (18.5%) enrolled in fifth grade. Of those participants enrolled in school, twenty-two (81.5%) received general education instruction, and four (14.8%) received special education instruction. Demographic information is represented in Table 1.

Twenty-five of the participants (92.6%) were born in the U.S., and two (7.4%) were born in Guatemala. Although most of the participants were American, their ethnic backgrounds varied. Specifically, six participants (22.2%) were Dominican, four (14.8%) were Ecuadorian, one (3.7%) was Ecuadorian/Dominican, one (3.7%) was Ecuadorian/Dominican/Haitian, two participants (7.4%) were Ecuadorian/Mexican, and one participant (3.7%) was Ecuadorian/Mexican/Italian. In addition, three participants (11.1%) were Guatemalan, six (22.2%) were Mexican, one (3.7%) was Mexican/ Colombian, one (3.7%) was Puerto Rican/Mexican, and another participant (3.7%) was Salvadorian. A visual representation of this information is available in Figure 1.

Language Demographics

Spanish was the native language of twenty-five participants (92.6%). One of the participants (3.7%) was a native Spanish and Q'eqchi' speaker. Another participant's (3.7%) native language was also a Mayan language, Mam, in addition to Spanish. Five participants (18.5%) were first exposed to English at one-year-old while another five

(18.5%) were exposed at two years old. Eight participants (29.6%) were exposed to English at three years old. Six participants (22.2%) were exposed to English at four years old, while two participants (7.4%) were first exposed at five years old, and one participant (3.7%) at six years old. The number of participants and their percentage of English language exposure is as follows: one participant (3.7%) had 91% English language exposure in their lifespan, one participant (3.7%) had 89% English exposure in their lifespan, one participant (3.7%) had 88%, three participants (11.1%) had 80%, two participants (7.4%) had 78%, one participant (3.7%) had 73%, two participants (7.4%) had 71%, one participant (3.7%) had 70%, two participants (7.4%) had 67%, one participant (3.7%) had 64%, one participant (3.7%) had 60%, three participants (11.1%) had 57%, three more (11.1%) had 50%, one participant (3.7%) had 44%, one participant (3.7%) had 43%, two participants (7.4%) had 33%, and one participant (3.7%) had 25% of English language exposure. Participant language demographic information is summarized in Table 2.

Measures

Two assessment measures were used in the present study to test the hypothesis. Both measures are described further in the proceeding section.

Ortiz Picture Vocabulary Acquisition Test (Ortiz PVAT)

The Ortiz PVAT is a receptive vocabulary test that was normed with two samples, ELs and ESs in the U.S., between the ages of 2 years 6 months and 22 years 11 months old. ELs were identified as individuals who were exposed to another language other than English. The EL sample group included 1,190 individuals and was nearly equally split by gender, with 49.2% male and 50.8% female. A stratified sample was used based on the

individual's ethnic and racial background, age, gender, their parents' level of education, geographic region, and length of English exposure. ESs were identified as individuals whose only home language was English. A total of 1,530 ESs were included in the sample group. Gender was equally divided (50% each – male/female) in the sample group. A stratified sample was used based on the individual's ethnic and racial background, age, gender, their parents' level of education, and geographic region; stratifying the sample allowed for representation of the population.

The Ortiz PVAT examines diverse word categories that range in difficulty considering various parts of speech (e.g., nouns, verbs, adjectives, adverbs, and prepositions) and BICS and CALPS words receptively understood. Words involving actions, feelings, foods, activities, math and science terminology, toys, animals, body parts, household items and such are included in the assessment. Within word categories, target words vary in range of difficulty based on word frequency and grade-level demands. Two forms, Form A and Form B, each include a total of 167 items. Every fifth item is presented in both forms in order to ensure comparability. Of the items selected for use in the test, consideration was made with regard to the frequency of word category, type (BICS and CALPS), part of speech, and the position of the target image.

The test is administered and scored via an iPad or tablet. Administration lasts approximately 15 to 30 minutes. The examiner presents the student with the iPad or tablet and delivers the task instructions as provided by the test developer. Four pictures are visually presented on the tablet, including the picture of a target word and three distracting images. The images are grouped in the formation of a quadrant (two images beside one another and placed above two more images). The target word is presented in

print (below the four pictures) and orally by the software. The student can listen to the target word as many times necessary. Then, the student uses their finger or stylus to click on the visual image of the target word. The software independently determines basals and ceilings (five errors within ten items presented consecutively) based on the individual's responses. Screener items assist with determining start points. Assessment and Progress Reports can be generated electronically.

Woodcock Muñoz Language Survey-Third Edition (WMLS III)

The WMLS III is a standardized assessment that evaluates a student's language proficiency and assists in guiding placement for ELs (HMH, 2019). The WMLS III was developed utilizing the norming sample of the Woodcock-Johnson Fourth Edition (WJ IV; Woodcock, Alvarado, Ruef, & Schrank, 2017). Specifically, 7,416 between the ages of 2 to over 90 years old participated in the norming process. This sample was comprised of ESs residing in diverse communities in the U.S. The sample breakdown is as follows: 664 children between the ages of 2 and 5 years old, not enrolled in kindergarten, 3,891 kindergarten to twelfth-grade students, 775 college/university students, and 2,086 adults. Random selection within a stratified sample allowed for control of examinee and community variables (e.g., ethnicity, sex; Woodcock, Alvarado, Ruef, & Schrank, 2017).

As the WMLS III was co-normed with the WJ IV, the age range of the population sample is large; however, only the norms from 3 years 0 months to 22 years 11 months old apply to the WMLS III (Woodcock, Alvarado, Ruef, & Schrank, 2017). A calibration study, using a few hundred new English and Spanish items, was conducted as part of the WMLS III process and included 1,055 native ESs and 1,041 native Spanish Speakers (SS). This data was used to calibrate the added test items and to ensure the items

corresponded with the tests' underlying scales (Woodcock, Alvarado, Ruef, & Schrank, 2017).

The WMLS III compares EL students to others of the same grade (HMH, 2019). Two English forms, Form A and Form B, and one Spanish form are available for administration. The English and Spanish forms are both comprised of eight tests – Test 1: Analogies/Prueba 1: Analogías, Test 2: Oral Comprehension/Prueba 2: Comprensión oral, Test 3: Picture Vocabulary/Prueba 3: Vocabulario sobre dibujos, Test 4: Oral Language Expression/Prueba 4: Expresión de lenguaje oral, Test 5: Letter-Word Identification/Prueba 5: Identificación de letras y palabras, Test 6: Passage Comprehension/Prueba 6: Comprensión de textos, Test 7: Dictation, Prueba 7: Dictado, and Test 8: Written Language Expression/Prueba 8: Expresión de lenguaje escrito (Woodcock, Alvarado, Ruef, & Schrank, 2017).

Tests can yield an individual score or, when grouped, can form clusters based on the task demands and provide cluster scores (Woodcock, Alvarado, Ruef, & Schrank, 2017). For example, administering and scoring Test 1: Analogies/Prueba 1: Analogías and Test 3: Picture Vocabulary/Prueba 3: Vocabulario sobre dibujos can yield independent scores for each or when combined provide the Basic English Oral Language cluster score or, in Spanish, Lenguaje oral básico en español, both measures of foundational speaking and listening skills in their respective language. The reliability of this cluster with students between the ages of 5 and 18 is .89 (Woodcock, Alvarado, Ruef, & Schrank, 2017).

The present study considered the performance of Spanish-speaking ELs on Test 1: Analogies and Test 3: Picture Vocabulary, both measures of lexical knowledge. Test 1

requires students to actively listen and comprehend the relationship between two words and apply that knowledge to determine and provide a single missing word to another pair of words. Test 3 requires the student to verbally identify pictured objects presented using single words. The reliability of Test 1 with students between the ages of 5 and 18 years old is .86, and the reliability of Test 3 with students between the ages of 5 and 18 years old is .79 (Woodcock, Alvarado, Ruef, & Schrank, 2017).

Procedure

Once contact with parents who expressed interest in the study was made, the researcher answered any questions the parents had and sought consent. Some parental consent forms were obtained electronically, while others were mailed. For those parents and children who were not available to participate at the moment, an appointment was scheduled. In the first session, the researcher introduced herself to the child/participant and sought their assent/approval to participate in the study. The child's demographic information and language background were also obtained. The form includes information from the WMLS III protocol and Ortiz PVAT student profile, as well as questions regarding the participant's race, ethnicity, and educational placement (e.g., general education or special education classification). At the end of the first session, the researcher scheduled a day and time to administer the assessments virtually via Zoom (Version 5.7.4). Parents of participants under the age of 8 and in grades prekindergarten to first grades were asked to utilize a tablet or iPad in order for the researcher to share remote control access for early items in Test 3: Picture Vocabulary.

Due to restrictions on face-to-face interactions resulting from the Coronavirus (COVID-19) pandemic, the Ortiz PVAT and WMLS III subtests were administered

virtually (also known as Remote Testing). Every effort was made to simulate standardized assessment practices as outlined within the manuals of each assessment tool. In addition, the researcher participated in professional development/webinars provided by both publishers as well as district-wide training in remote testing. The publishers also offered considerations and tips when conducting remote testing. Under the ethical standards of NASP, remote testing is permissible as long as the examiner possesses the skills to deliver the assessments in person and the tools used remotely are appropriately adapted for this method of delivery.

The research tasks were scheduled during summer break, around students' school activities (for those in summer school or camp), and administered during the second/final session. The first session lasted approximately 15 minutes, and the final session lasted approximately 30 minutes. With younger participants as well as participants who struggled to sustain attention, the administration of tasks varied from 40 minutes to 50 minutes. Parents of younger students were asked to remain near the student to assist with redirection if necessary. No audio or videotaping occurred. In order to minimize priming effects, assessments were delivered in a counterbalance format. With the three assessments (WMLS III Test 1: Analogies, Test 3: Picture Vocabulary, and Ortiz PVAT) to be utilized in the study, there are six possible administration sequences/orders. The researcher assigned a number to each set of possible administration sequences and used a random generator (Urbaniak & Plous, 2021) to randomly predetermine the sequence of assessment administration for each potential participant.

The researcher encouraged participants, as indicated in assessment manuals and as the researcher determined necessary considering the participant's motivation. Upon

completing all vocabulary tests administered, all students were offered small prizes (e.g., sharpener, pencils, pencil grip, erasers, stickers, coloring book). Prizes were mailed at the end of the week. In addition, parents were mailed a $10 gift card. Results from the Ortiz PVAT and the WMLS III Test 1: Analogies and Test 3: Picture Vocabulary were scored and generated using their respective online scoring system.

Data Analysis

The researcher developed a Microsoft Excel spreadsheet maintaining each student's demographic information (e.g., date of birth, gender, grade, amount of English exposure, general education, or special education classification) alongside a pseudonym (i.e., AG-001). The researcher summarized the sample characteristics from the raw data, including the demographic variables as well as the outcome data. Categorical variables were expressed as counts and percentages. Continuous variables were expressed as means and Standard Deviations (SD). Variables were examined for appropriate distributions, particularly the continuous variables, which were examined for an appropriate range, outliers, and normality. For the purpose of this study, a standard score of 40 for the WMLS III and Ortiz PVAT was the minimum score allowed.

The researcher first conducted a Pearson correlation coefficient to examine the direction and degree of the relationship between the Ortiz PVAT and the WMLS III English. Next, One-Sample T-Tests were conducted to compare the study sample means of the Ortiz PVAT and WMLS III against the population means, as outlined respectively in each testing manual. The Ortiz PVAT and WMLS III both have a standard score mean of 100 and SD of 15. Additionally, Dependent or Paired Samples T-Tests were conducted to determine whether there was a statistically significant difference between the score

means of the Ortiz PVAT and WMLS III English. Specifically, three comparisons were made utilizing four scores: the standard scores on the Ortiz PVAT were compared against the standard scores from Test 1: Analogies, Test 3: Picture Vocabulary, and the Basic English Oral Language Cluster from the WLMS III English form. Cohen's *d e*ffect sizes were also computed.

A moderate effect size was expected given the differences in the norming structures between the Ortiz PVAT and WMLS III, where the former includes bilingual, exposure-based norms, and the latter does not. A one-tailed test was utilized, as the researcher hypothesized that Spanish-speaking ELs would perform significantly differently on both assessments with a deviation below the normative average likely on the WMLS III scores. A priori power analysis using G*Power 3.1.9.5 (Faul et al., 2009) indicated that to achieve a power of 0.80, as recommended by Cohen (1992), with $p <$ 0.05 and a moderate effect size (0.6), 19 participants would be required.

Individual standard score performance was compared graphically to observe any potential trends. Additionally, the relationship between Spanish-English learners' English language exposure and their performance in the Ortiz PVAT and WMLS III English – Test 1: Analogies, Test 3: Picture Vocabulary, and Basic English Oral Language Cluster was compared graphically. Finally, further comparisons were made between individual Ortiz PVAT and WMLS III Test 3: Picture Vocabulary standard scores and between these standard scores and the participant's English language exposure. As mentioned, sample size analysis was conducted with G*Power 3.1.9.5. All other statistical analyses were performed using IBM SPSS Software (Version 28). All statistical tests were one-tailed with alpha = 0.05.

Results

A Pearson correlation coefficient was computed to assess the relationship between the Ortiz PVAT and WMLS III English – Test 1: Analogies, Test 3: Picture Vocabulary, and the Basic English Oral Language Cluster. There was a moderate positive correlation between the Ortiz PVAT and WMLS III Test 1: Analogies, $r = +.506$, $n = 27$, $p = .004$. There was a strong positive correlation between the Ortiz PVAT and WMLS III Test 3: Picture Vocabulary, $r = +.709$, $n = 27$, $p = <.001$. There was a strong positive correlation between the Ortiz PVAT and WMLS III Basic English Oral Language Cluster, $r = +.720$, $n = 27$, $p = <.001$. Results of the Pearson correlations are summarized in Table 4.

The expected sample average using the Ortiz PVAT and WMLS III was 100. Both measures have a SD of 15. The mean score obtained on the Ortiz PVAT (M = 107.9, SD = 6.9) was 7.9 points larger than the expected sample average of 100. The Ortiz PVAT mean score was 16.7 points higher than the mean score obtained on the WMLS III Test 1: Analogies (M = 91.2, SD = 14.2), 24.0 points higher than the mean score obtained on the WMLS III Test 3: Picture Vocabulary (M = 83.9, SD = 15.0), and 23.3 points higher than the WMLS III Basic English Oral Language Cluster (M = 84.6, SD = 15.0). Results are visually represented in Table 6 and graphically represented in Figure 2.

Results of the One-Sample T-Test determined that the mean score obtained on the Ortiz PVAT (M = 107.9, SD = 6.9) was significantly higher than the population mean $t(26) = 5.94$, 95% Confidence Interval of the Difference [5.14, 10.57], $p < .001$). The WMLS III Test 1: Analogies mean score (M = 91.2, SD = 14.2) was significantly lower than the population mean $t(26) = -3.22$, 95% Confidence Interval of the Difference [-

8.82, -14.44], p = .002). The WMLS III Test 3: Picture Vocabulary mean score (M = 83.9, SD = 15.0) was significantly lower than the population mean $t(26)$ = -5.58, 95% Confidence Interval of the Difference [-16.07, -21.99], $p < .001$). The WMLS III Basic English Oral Language Cluster (M = 84.6, SD = 15.0) was also significantly lower than the population mean $t(26)$ = -5.33, 95% Confidence Interval of the Difference [-15.37, -21.30], $p < .001$). A summary of these results is provided in Table 5.

A Cohen's *d* analysis of the One-Sample T-Tests demonstrated a large effect size for WMLS III Test 1: Analogies (-0.62), and exceptionally large effect sizes for the Ortiz PVAT (1.14), WMLS III Test 3: Picture Vocabulary (-1.07), and WMLS III Basic English Oral Language Cluster (-1.03). The effect sizes are summarized in Table 5 and illustrated in Figure 3.

Results of the Dependent or Paired Samples T-Test between the Ortiz PVAT and the WMLS III Test 1: Analogies were statistically significant ($t(26)$ = 7.07, 95% Confidence Interval of the Difference [11.82, 21.52], $p < .001$). In addition, the results of the Paired Samples T-Test between the Ortiz PVAT and the WMLS III Test 3: Picture Vocabulary were statistically significant ($t(26)$ = 11.10, 95% Confidence Interval of the Difference [19.50, 28.36], $p < .001$). Lastly, the results of the Paired Samples T-Test between the Ortiz PVAT and the WMLS III Basic English Oral Language Cluster were also statistically significant ($t(26)$ = 10.84, 95% Confidence Interval of the Difference [18.82, 27.63], $p < .001$). A summary of these results is provided in Table 6.

Not only did the Paired Samples T-Test yield statistically significant results between the Ortiz PVAT and the WMLS III English, but a Cohen's *d* analysis demonstrated exceptionally large effect sizes for all three comparisons. The Ortiz PVAT

and WMLS III Test 1: Analogies yielded an effect size equivalent to 1.36. The Ortiz PVAT and WMLS III Test 3: Picture Vocabulary yielded an effect size equivalent to 2.14. The Ortiz PVAT and WMLS III Basic English Oral Language Cluster yielded an effect size equivalent to 2.09. Effect sizes are summarized in Table 6 and illustrated in Figure 4.

The individual standard score performance of Spanish-English learners (ages 5 – 11) across the Ortiz PVAT and WMLS III English – Test 1: Analogies, Test 3: Picture Vocabulary, and Basic English Oral Language Cluster are graphically illustrated in Figure 5. Additionally, a comparison of the amount (percent) of English language exposure of Spanish-English learners (ages 5 – 11) and their individual standard score performance in the Ortiz PVAT and WMLS III English – Test 1: Analogies, Test 3: Picture Vocabulary, and Basic English Oral Language Cluster are graphically illustrated in Figure 7; the amount of English exposure was organized from least to greatest to demonstrate trends. In cases where the participants had the same amount of English exposure, the data was organized by age (youngest to oldest). Of all the tasks, the WMLS III Test 3: Picture Vocabulary is most similar to the Ortiz PVAT; as such, direct comparisons between individual Ortiz PVAT and WMLS III Test 3: Picture Vocabulary standard scores are made and illustrated in Figure 6. A comparison of these standard scores against the participants' English language exposure is illustrated in Figure 8.

A visual review of the graphs indicates that the Ortiz PVAT is consistently higher than the standard scores on the WMLS III; this is most evident in comparing the Ortiz PVAT and WMLS III Test 3: Picture Vocabulary (Figures 6 and 8). The lowest obtained standard score in the Ortiz PVAT was 93, while the lowest standard score in the WMLS

III Test 3: Picture Vocabulary was 48. The lowest standard score earned in the WMLS III Test 1: Analogies was 65. The lowest WMLS III Basic English Oral Language Cluster standard score was 50. Table 3 provides the minimum and maximum scores obtained for each of the tests and cluster. Of the 27 participants included in the study, 13 participants performed below the average range (Standard Score < 90) in the WMLS III Test 1: Analogies, 18 participants performed below the average range in the WMLS III Test 3: Picture Vocabulary, and 14 participants had below average scores in the WMLS III Basic English Oral Language Cluster. Furthermore, seven of the 18 participants who performed below average in the WMLS III Test 3: Picture Vocabulary had over 70% English exposure across their life. No participant performed below the average range in the Ortiz PVAT.

Discussion

As noted in the literature review, vocabulary development is essential in academic success. ELs are challenged with learning the BICS of a new language while also learning academic language (CALPS). This leaves this population at a disadvantage from their same-grade ES peers who have established BICS in English and are prepared to learn academic language. ELs perform significantly below same grade peers, are overrepresented in special education, and often drop out of school compared to their ES peers.

As a field, assessments and appropriate evaluation methods are necessary to make informed decisions about the education of ELs. When assessing achievement and developing instructional interventions, school psychologists and educators must consider the impact of English language acquisition and proficiency (Rabinowitz, 2008). There has been insufficient research conducted studying the language and literacy development of ELs (Hammer et al., 2014). Specifically, as Ortiz (2019) noted, there has been a lack of research on language acquisition as it pertains to the development of assessments outside of native-language and nonverbal assessments. "…Our present understanding of EL test performance is more intuitive than empirical" (Ortiz, 2019).

ELs must be compared to others who resemble them most. As most language proficiency tools fail to include appropriate true peer norms, several factors should be considered when evaluating language proficiency and abilities. As explained by Collier (1989), the age of the EL when he/she arrives in the U.S., as well as his/her length of exposure to English, is essential in English language acquisition and, ultimately, proficiency. Other important factors include the history of the student's formal education

in their L1, proficiency of L1, the amount of time the student can practice L2 with peers who speak the language fluently, and the student's BICS and CALPS in both languages (Elizalde-Utnick & Romero, 2017). As of present, the majority of assessment tools utilized with Spanish-speaking ELs in the U.S. were normed with monolingual Spanish students from other countries. When considering the challenges ELs experience with regard to language acquisition and proficiency, acculturation, and overall English academic instruction, among other societal challenges (i.e., low socioeconomic status), it is disturbing to see such large numbers of ELs deemed eligible for special education services and dropping out of school.

The purpose of the present study was to determine whether there is a difference in the vocabulary knowledge of Spanish-speaking ELs when controlling for developmental language norms. This researcher hypothesized the following: assuming the WMLS III's norming sample is potentially discriminatory because it is based on monolingual, non-exposure norms, Spanish-speaking ELs would perform significantly differently (deviation below the normative average is likely) on the WMLS III than on the Ortiz PVAT due to the latter's use of bilingual, exposure-based norms. Additional goals of this study included identifying equitable assessment measures for ELs that can be utilized to make data-driven decisions with regard to the unique educational needs of ELs and special education referrals, increasing assessment knowledge and literature of ELs, as well as improving the field's insight of developmental language norms in ELs, which will further allow school psychologists to enhance their skills in working with this population.

Correlational analyses between the Ortiz PVAT and the WMLS III English determined strong, positive relationships among the tests. The findings are similar to

Adamek (2019), who evaluated language abilities among ELs and determined that the Ortiz PVAT and the Verbal Comprehension Index (VCI) in the Wechsler Intelligence Scales for Children – Fifth Edition (WISC-V) had strong correlations, particularly with ELs residing in the U.S. Of all three correlations in the present study, the Ortiz PVAT and WMLS III Basic English Oral Language Cluster yielded the strongest relationship (r = +.720); however, the Basic English Oral Language Cluster is a score developed from the grouping of the WMLS III - Test 1: Analogies and Test 3: Picture Vocabulary scores. When reviewing the correlational analysis between the Ortiz PVAT and each of the subtests that comprise the Basic English Oral Language Cluster, the Ortiz PVAT and the WMLS III Test 3: Picture Vocabulary also had a strong positive correlation (r = +.709). Language accounts for approximately 50% of the correlation between the Ortiz PVAT and Test 3: Picture Vocabulary. Of the three tests, Test 3: Picture Vocabulary is most similar to the Ortiz PVAT; both tasks present images to the student and require the student to demonstrate an understanding of the visual representation(s). The Ortiz PVAT and WMLS III Test 1: Analogies had the weakest, although moderate, correlation (r = +.506). Language accounts for approximately 25% of the correlation between the Ortiz PVAT and Test 1: Analogies; this could be due to several reasons. First, Test 1: Analogies measures expressive language while the Ortiz PVAT measures receptive language. Second, task demands in Test 1: Analogies are not solely dependent on language-based knowledge (*Gc*) but also incorporate some degree of cognitive ability, specifically Fluid Reasoning (*Gf*). *Gf* "is the ability to reason, form concepts, and solve problems using unfamiliar information or novel procedures" (Schrank et al., 2010). These differences could explain the moderate correlation between the Ortiz PVAT and Test 1:

Analogies. Adamek (2019) found similar results suggesting that the performance of ELs on the Similarities (SI) subtest in the WISC-V was impacted by *Gf*.

A One-Sample T-Test determined that the difference between the Ortiz PVAT and the WMLS III English is significantly large, with the population mean falling between both measures. The mean standard score obtained on the Ortiz PVAT (M = 107.9, SD = 6.9) was significantly higher than the population mean, where the mean scores of the WMLS III Test 1: Analogies (M = 91.2, SD = 14.2), WMLS III Test 3: Picture Vocabulary (M = 83.9, SD = 15.0), and WMLS III Basic English Oral Language Cluster (M = 84.6, SD = 15.0) were significantly lower than the population mean. When looking closely at the results between the WMLS III scores, Test 1: Analogies appears to be – to some extent – fairer in assessing ELs compared to Test 3: Picture Vocabulary and Basic English Oral Language Cluster. This is evident particularly as the latter two both fall within the end of the Low Average range, and Test 1: Analogies falls in the Average range. Another notable finding was that the Ortiz PVAT standard score mean was nearly eight points above the expected standard score average of 100. Adamek (2019) made a similar revelation in her study when she obtained a mean score difference of approximately eight points above average on the Ortiz PVAT.

It is unclear what variables caused the Ortiz PVAT mean scores to be several points higher than the normative average in both studies; however, ELs still performed within the average range (Standard Score: 90 – 110) in the Ortiz PVAT. Therefore, it would be unlikely to misidentify an EL as having language delays using this test with exposure-based norms. In contrast, ELs tended to earn standard scores below the average range (Standard Score < 90) in the WMLS III. If the mean score on the Ortiz PVAT were

closer to the expected average, the differences between the means would remain large, with the Ortiz PVAT being approximately 16 points higher than the WMLS III Test 3: Picture Vocabulary. Also, if the scores on the Ortiz PVAT were lowered 7.9 points to the population average of 100 and the same number of points were reduced on the WMLS III tests, then the mean scores would dramatically change the descriptive ranges from Average to Low Average in Test 1: Analogies, and from Low Average to Low in Test 3: Picture Vocabulary and Basic English Oral Language Cluster. This would identify ELs as further delayed in vocabulary knowledge when assessed with this tool.

It is essential to consider that although all three tests measure lexical knowledge, there are differences in task demands. Specifically, the Ortiz PVAT is purely receptive, while Test 3: Picture Vocabulary initially measures receptive language with later items placing expressive language demands on students. Receptive skills are developed before expressive skills (North Shore Pediatric Therapy, 2021) and are generally more advanced in typically developing children (Seol et al., 2014). Regardless of the differences in task demands, participants should yield average scores in the Ortiz PVAT and WMLS III, based on normative means outlined in each technical manual; however, this was not reflected in this study.

Results of the One-Sample T-Tests yielded statistically significant values. Moreover, exceptionally large effect sizes were found in the One-Sample T-Tests, particularly with the Ortiz PVAT (1.14), WMLS III Test 3: Picture Vocabulary (-1.07), and WMLS III Basic English Oral Language Cluster (-1.03) as compared to the population mean. The WMLS III Test 1: Analogies (-0.62) had a large effect size. The

Ortiz PVAT and population mean held the strongest relationship, indicating that this assessment tool yielded scores closest to the normative value.

Overall, results of the One-Sample T-Test comparing the Ortiz PVAT with the population mean suggests that Spanish-speaking ELs are learning English at a rate that is typical (within normal limits) for their age and amount of English language exposure. Even if the Ortiz PVAT standard score mean was overestimated, it does not risk wrongfully classifying students with an educational disability. In contrast, the use of the WMLS III routinely underestimates ELs lexical knowledge and places students at risk from misidentification in special education.

Statistically significant results (p <.001) were also found in the Paired Samples T-Test between the Ortiz PVAT and WMLS III Test 1: Analogies, Test 3: Picture Vocabulary, and Basic English Oral Language Cluster. It is highly unlikely that the difference occurred solely by chance. Furthermore, significantly large effect sizes were calculated between the Ortiz PVAT and WMLS III Test 3: Picture Vocabulary (2.14), the Ortiz PVAT and Basic English Oral Language Cluster (2.09), and the Ortiz PVAT and WMLS III Test 1: Analogies (1.36). These results (Table 6) indicate that the Ortiz PVAT, which utilizes exposure-based norms, yielded notable differences in measuring lexical knowledge compared to the WMLS III, which does not consider the language development of ELs.

The outcomes of all statistical analyses provide evidence in support of the hypothesis. The Paired Samples T-Test between the Ortiz PVAT and the WMLS III Test 1: Analogies, Test 3: Picture Vocabulary, and Basic English Oral Language Cluster yielded statistically significant results of p <.001. The Ortiz PVAT and WMLS III Test 3:

Picture Vocabulary held the largest effect size (2.14), and the Ortiz PVAT and Basic English Oral Language Cluster followed with an effect size of 2.09. The Ortiz PVAT and WMLS III Test 1: Analogies yielded the weakest effect size (1.36), although still large. Adamek (2019) also found moderate to large effect sizes between the Ortiz PVAT and the WISC-V VCI and between the Ortiz PVAT and the WISC-V VC/SI [Vocabulary/ Similarities]. A study conducted by Tello (2020) evaluating ELs classified with SLI using tools with bilingual, exposure-based norms (i.e., Ortiz PVAT) and tools lacking these norms (i.e., WISC-V VCI and Clinical Evaluation Language Fundamentals-5 [CELF-5]) determined even larger effect sizes (Cohen's *d*) ranging from 2.01 to 3.77. Furthermore, results of the One-Sample T-Tests in the present study found that ELs are at greater risk of being misidentified in special education when using the WMLS III, given the large underestimates in their performance. In combination with prior studies, the present study, suggests that assessments developed lacking bilingual, exposure-based norms such as the WMLS III, WISC-V, CELF-5 and such are inappropriate to use with ELs in the U.S.

In addition to statistical tests, the individual performance of ELs was visually displayed across several graphs and reviewed for trends. When reviewing the trend lines of Figure 7, the scores of three participants on Test 1: Analogies surpassed their scores on the Ortiz PVAT, Test 3: Picture Vocabulary, and Basic English Oral Language Cluster. As noted earlier, solving for analogies involves cognitive skills such as inferential reasoning (*Gf*) that are not required in the other presented tasks. These three participants likely have fluid reasoning skills (*Gf*) that are better developed than their language-based knowledge/crystallized intelligence (*Gc*). Average and above-average scores cannot be accidentally obtained.

Adamek (2019) reported that it is likely for ELs assessed with the WISC-V to perform in the Low range while performing in the Average range on the Ortiz PVAT. In her study, she determined that in nearly all cases, "score[s] on the WISC-V VCI and WISC-V VC/SI subtests were lower, often statistically significant so, than scores on the Ortiz PVAT" (Adamek, 2019, p. 35). When considering the participants' lifetime exposure to English (Figures 7 and 8), even participants with low amounts of English exposure performed in the average range in the Ortiz PVAT; this is due to the exposure-based norms in the Ortiz PVAT. Moreover, almost 40% of the 18 participants who earned average scores in the Ortiz PVAT but below-average scores in Test 3: Picture Vocabulary had over 70% lifetime exposure to English. This provides further evidence demonstrating the value of considering the language development of ELs when developing assessments for this population. As reported earlier, the WMLS III does not consider exposure norms and, ultimately, language development of ELs.

In the study conducted by Mancilla-Martinez and Lesaux (2011), researchers utilized the Woodcock Language Proficiency Battery-Revised to assess the vocabulary growth of ELs. The Spanish form was adapted from the English form and was normed using monolingual populations. Although the researchers indicate that "each form contains unique item content, allowing scores from the two tests to be compared without concerns," they later determined that all three groups (mostly Spanish, equal amounts of Spanish/English, and mostly English language exposure/use) performed "well below national norms in Spanish" with the latter being outperformed (Mancilla-Martinez & Lesaux, 2011, pp. 538, 541). This information is indicative that monolingual norms, be that Spanish or English, are not valid measures with bilingual students. Monolingual SSs

perform well in Spanish and monolingual ESs perform well in English, while ELs perform comparatively worse than the other two populations. The WMLS III's norming sample is potentially discriminatory because it is based on monolingual, non-exposure norms. Assessments like the WMLS III continue to be developed and remain most unjust for ELs.

In an article written by Ortiz and Wong (2020), they discuss WMLS III data of Spanish-speaking ELs provided by a suburban school district in the Southeastern part of the U.S. Of the 14 students assessed using the WMLS III, 12 students performed "well below normal limits" and were at risk of being classified with an educational disability (Ortiz & Wong, 2020). In comparison, when using the Ortiz PVAT, only one of these same students performed in the Very Low range and potentially had an educational disability (Ortiz & Wong, 2020). These results coincide with those of the current study and are alarming considering the fact that the WMLS III, as well as other tools with similar norming structures, are utilized in school districts across the country to assist in making educational determinations for ELs.

All findings in this study show a substantial difference between the scores of ELs in the Ortiz PVAT and the WMLS III English in support of the hypothesis. This study determined that the use of assessments with bilingual, exposure-based norms such as the Ortiz PVAT produced higher scores in comparison to assessments using monolingual, non-exposure norms when given to ELs, and the difference was statistically significant. A study conducted by Tello (2020) also obtained similar findings. The present study found exceptionally large effect sizes and differences between the means for all three comparisons (Ortiz PVAT and WMLS III Test 1: Analogies, Test 3: Picture Vocabulary,

and the Basic English Oral Language Cluster). Tello (2020) also found significantly large effect sizes and differences between the means when comparing the Ortiz PVAT to a number of language-based tests with non-exposure norms. With the Ortiz PVAT consistently yielding scores in the average range while the WMLS III English subtests yielded many below average scores, it can be presumed that the norming sample of the WMLS III is discriminatory towards ELs and problematic when utilized to determine the language abilities of ELs or answer special education referral questions.

Limitations

When the researcher initially obtained IRB approval, the study was to be conducted within a school setting. Only one school district approved the study, and only one school within the district agreed to allow the study to be conducted with their students. There were notable challenges with obtaining district approval, given the study was being initiated in the new school year (2020-2021) after the Coronavirus global pandemic (also known as COVID-19) forced schools to close and provide remote instruction during the middle of the prior school year (March 2020). School leaders, administrators, educators, clinicians, students, and their families around the world were coping with unique, daily challenges never before seen in their lifetime. The study was open to the aforementioned school for three months; due to low recruitment, it was necessary to change the site of the study from the school to the general community to maximize the sample size. Although this study was not conducted in a school setting, students in the target age range are school-aged children, and the results can be generalized to the population.

A limitation of this study was the small sample size ($N = 27$); some Spanish-speaking parents were hesitant to have their children participate in the study, possibly due to fear that their immigration status would be exposed. Despite the researcher's efforts at explaining confidentiality, some families did not log onto the scheduled Zoom session or answer follow-up phone calls; this was a factor that could not be controlled.

Another limitation of this study is that the assessments were administered to participants virtually/remotely; however, both tools were normed in-person/face-to-face. Typically, only the student and examiner work together in the session area with in-person assessments. Remote administration in the present study required the assistance of parents to redirect younger children as well as those who were distractible. This method of assessment was utilized due to restrictions on in-person evaluations resulting from the Coronavirus (COVID-19) global pandemic. In order to enhance skills in remote testing, the researcher participated in several webinars, including those provided by the author of the Ortiz PVAT and the publishers of the WMLS III, Riverside Insights. Additionally, the researcher engaged in remote practice sessions with several colleagues using both tools.

Although standard scores obtained in this assessment need to be interpreted with caution due to a change in the administration of the assessments from that of the norming sample, more accurate scores would not be yielded from face-to-face testing at this time. Given the physical, social, psychological, and emotional changes individuals may have undergone from the Coronavirus, the norms in assessments developed prior to the pandemic no longer apply to that of the current population, which makes comparisons to the norming sample questionable. In other words, in-person assessments do not guarantee accurate, valid comparability. Remote testing was the most appropriate method of

assessment at the time of this study. Despite the use of remote assessment, it is crucial to consider that if this administration method impacted performance, it would have affected all testing outcomes, not simply one or any test randomly. For example, participants would have consistently obtained scores in the extremely high/superior range and/or extremely low range across both measures. Remote administration did not appear to violate standardization and evidently did not impact participants' performance.

One limitation of the study that is not a direct result of the Coronavirus (COVID-19) global pandemic is that each tool utilized in the present study measure lexical knowledge with variations in each. Test 1: Analogies is orally presented to the examinee, while Test 3: Picture Vocabulary and the Ortiz PVAT present visuals. Early items on Test 3: Picture Vocabulary begin with a receptive vocabulary component and transition to expressive vocabulary using animated illustrations. The Ortiz PVAT measures vocabulary in a purely receptive format using real-life images. With this in consideration, ELs still tended to perform below average in the WMLS III compared to the normative sample outlined within the technical manual.

Implications for Practice and Directions for Future Research

Considering the different normative standards in the Ortiz PVAT and WMLS III English, this study evaluated the vocabulary knowledge of Spanish-speaking ELs using each tool. The results proved to be beneficial in expanding research and theory. Specifically, the results demonstrated that using tools that incorporate exposure-based norms (e.g., Ortiz PVAT) offers a more accurate measure of vocabulary knowledge as observed through performance scores falling in the average range as well as large effect sizes when measured against tests with non-exposure-based norms. The Ortiz PVAT can

be utilized to assist in making data-driven decisions concerning the unique educational needs of ELs and special education referrals. The findings ultimately provide added support to the need for true peer comparisons. Use of assessments like the WMLS III that fail to consider critical variables (i.e., amount of English language exposure) when developed will continue contributing to the inappropriate identification of ELs as having an educational disability. Relying on monolingual norms does not work. The use of true peer comparisons (i.e., exposure-based norms) represents the next evolution in evaluating ELs, and test publishers and authors need to include exposure-based norms when developing future assessments and tools.

Nationally certified school psychologists are responsible for advocating for culturally and linguistically appropriate tools and practices for students and families. Information obtained from this study has increased assessment knowledge and literature of ELs, particularly remote assessment, and improved the field's insight of developmental language norms in ELs. This research will further allow school psychologists to enhance their skills in working with this population. Additionally, this study has demonstrated that the use of remote assessment can yield accurate scores so long as the technology has been adequately vetted and steps delineated to address barriers prior to the administration of the assessment or study. This study demonstrates promising outcomes that could change the tools, methods, and strategies utilized to evaluate culturally and linguistically diverse students. Future studies should expand on developing tools with bilingual, exposure-based norms that evaluate other abilities, including expressive (oral and written) language abilities of ELs.

The present study included a small sample size ($N = 27$) of Spanish-speaking ELs. The demographic variables of participants in this study provide further evidence of the growing diversity in the U.S. Future studies should include larger samples as well as the participation of a greater number of children who recently arrived in the U.S. and have limited exposure to English. Although it is not recommended to evaluate children recently arriving in a new country, studies involving recently immigrated children can provide a better understanding of their lexical knowledge and further insight into the typical language development of ELs. Additionally, future studies should look at the lexical knowledge of multilingual learners exposed to a third language or dialect. The present study involved two participants whose native languages included Spanish and Q'eqchi' and another whose native languages included Spanish and Mam. Q'eqchi' is a Mayan language spoken by over one million people in Guatemala and Belize (Endangered Languages Project, n.d.). Mam, another Mayan language, is spoken by over half a million people across parts of Guatemala and Mexico (Endangered Languages Project, n.d.). One of the participants whose native languages included Spanish and Q'eqchi' recently immigrated to the U.S., and his performance presented as an outlier across all tests. This participant's data was removed, and analyses were not conducted. Had there been a larger sample size, including more participants who recently (within several months) arrived in the U.S. and spoke another language or dialect in addition to Spanish, the present research could have provided more insight on normal language development for these populations. Future research with populations who also speak indigenous languages and dialects is imperative given the large number of people immigrating to the U.S. from countries such as Guatemala, Peru, and Mexico where

indigenous languages (e.g., Q'eqchi', Quechua, Náhuatl, Yucatec Maya, Mam, and Mixteco) are spoken by over half a million people. Federal law requires schools to provide Free and Appropriate Public Education (FAPE) for all children (IDEA, 2004). The fields of education and school psychology need to continue building on the knowledge base, tools, and strategies used when working with culturally and linguistically diverse students.

As noted earlier, the present study "broke" standardization by administering the assessments remotely. This was conducted in this manner given the ongoing Coronavirus (COVID-19) global pandemic. Although the Coronavirus has impacted the population's characteristics in ways that are unclear at this time, future studies should consider administering the tasks in-person in the manner that the tools were previously normed – once restrictions are lifted, and Personal Protective Equipment (PPE) is no longer required. Future studies should also consider the potential differences in assessing students remotely versus in-person. Participants should be broken up into experimental (e.g., remote assessment only) and control (e.g., in-person assessment only) groups. Future tools used to evaluate the cognitive, academic, and linguistic abilities of all students should be normed/developed remotely and in-person, given the significant changes in learning and testing resulting from the Coronavirus global pandemic.

Conclusion

In summary, the results of this study supported the hypothesis; Spanish-speaking ELs performed significantly differently on the WMLS III than on the Ortiz PVAT, with a deviation below the normative average on the WMLS III scores. Spanish-speaking ELs performed significantly higher on the Ortiz PVAT than on the WMLS III due to its use of

bilingual, exposure-based norms, and the difference was statistically significant. The WMLS III norming sample is potentially discriminatory because it is based on monolingual, non-exposure norms. Although the intentions and ultimate goal of language proficiency measures such as the WMLS III are meaningful, these tools fail to incorporate the population of ELs residing in the U.S. and, albeit unintentionally, risk misidentifying ELs with an educational disability. The Ortiz PVAT is a promising tool; it allows for a valid and reliable comparison of vocabulary knowledge of ELs living in the U.S. who have more opportunities for English exposure than those living in their ethnic/native countries. Further research of ELs in the U.S., while considering the individual's age, amount of time exposed to English, and educational background, are necessary to expand on the field's knowledge and development of tools with these populations.

Table 1

Participant Demographics (N = 27)

Characteristic	*n*	%
Gender		
Male	12	44.4
Female	15	55.6
Age (years)		
5	2	7.4
6	3	11.1
7	6	22.2
8	2	7.4
9	7	25.9
10	4	14.8
11	3	11.1
Grade		
Not Enrolled	1	3.7
Prekindergarten	1	3.7
Kindergarten	2	7.4
1	6	22.2
2	2	7.4
3	5	18.5
4	5	18.5
5	5	18.5
Educational Placement		
General Education	22	81.5
Special Education	4	14.8
Not Enrolled	1	3.7
Country of Birth		
Guatemala	2	7.4
USA	25	92.6
Ethnicity		
Dominican	6	22.2
Ecuadorian	4	14.8
Ecuadorian/Dominican	1	3.7
Ecuadorian/Dominican/Haitian	1	3.7
Ecuadorian/Mexican	2	7.4
Ecuadorian/Mexican/Italian	1	3.7
Guatemalan	3	11.1
Mexican	6	22.2
Mexican/Colombian	1	3.7
Puerto Rican/Mexican	1	3.7
Salvadorian	1	3.7

Table 2

Language Demographics of Participants

Characteristic	*n*	%
Native Language(s)		
Spanish	25	92.6
Spanish/Mam	1	3.7
Spanish/Q'eqchi'	1	3.7
Age at First Exposure to English		
1 year old	5	18.5
2 years old	5	18.5
3 years old	8	29.6
4 years old	6	22.2
5 years old	2	7.4
6 years old	1	3.7
Percent of English Language Exposure Across Life		
0% to 25%	1	3.7
26% to 50%	7	25.9
51% to 75%	11	40.7
76% to 100%	8	29.6

Table 3

Descriptive Statistics for Ortiz PVAT and WMLS III English Standard Scores

Test (Standard Scores)	*N*	*M*	*SD*	Min	Max	*SEM*
Ortiz PVAT	27	107.9	6.9	93	119	1.32
WMLS III – Test 1: Analogies	27	91.2	14.2	65	122	2.73
WMLS III – Test 3: Picture Vocabulary	27	83.9	15.0	48	108	2.88
WMLS III – Basic English Oral Language Cluster	27	84.6	15.0	50	113	2.89

Table 4

Pearson Correlations for Ortiz PVAT and WMLS III English Standard Score Means

Test	*M*	*SD*	1	2	3	4
1. Ortiz PVAT	107.9	6.9	-			
2. WMLS III-Test 1: Analogies	91.2	14.2	.506**	-		
3. WMLS III-Test 3: Picture Vocabulary	83.9	15.0	.709**	.487**	-	
4. WMLS III-Basic English Oral Language Cluster	84.6	15.0	.720**	.841**	.880**	-

Note. **p significant at 0.01 (one-tailed); $N = 27$

Table 5

One-Sample T-Test Comparisons and Cohen's d Effect Sizes of Ortiz PVAT and WMLS III English Means against Normative Means

Test	*M*	*M* Difference	*t*	*df*	*p*	Cohen's *d*
Ortiz PVAT	107.9	7.9	5.94	26	<.001***	1.14
WMLS III-Test 1: Analogies	91.2	-8.8	-3.22	26	.002**	-0.62
WMLS III-Test 3: Picture Vocabulary	83.9	-16.1	-5.58	26	<.001***	-1.07
WMLS III-Basic English Oral Language Cluster	84.6	-15.4	-5.33	26	<.001***	-1.03

Note. Normative population mean = 100. *M* = Mean; *M* Difference = Mean Difference of standard scores; *t* = *t*-value; *df* = Degrees of Freedom; *p* = *p*-value.

***p* < .01 (one-tailed). ****p* < .001 (one-tailed).

Table 6

Paired Samples T-Tests and Cohen's d Effect Sizes for Ortiz PVAT and WMLS III English Standard Score Means

Test	*M*	*M* Difference	*SD*	*SEM*	*t*	*df*	*p*	Cohen's *d*
Ortiz PVAT – WMLS III-Test 1: Analogies	107.9 91.2	16.7	12.3	2.36	7.07	26	<.001***	1.36
Ortiz PVAT – WMLS III-Test 3: Picture Vocabulary	107.9 83.9	24.0	11.2	2.16	11.10	26	<.001***	2.14
Ortiz PVAT – WMLS III-Basic English Oral Language Cluster	107.9 84.6	23.3	11.1	2.14	10.84	26	<.001***	2.09

Note. Table demonstrates paired differences of *M* = Mean; *M* Difference = Mean Difference of standard scores; *SD* = Standard Deviation; *SEM* = Standard Error of the Mean; *t* = *t*-value; *df* = Degrees of Freedom; *p* = *p*-value.

****p* significant at .001 (one-tailed).

Figure 1

Participant Ethnicities (N = 27)

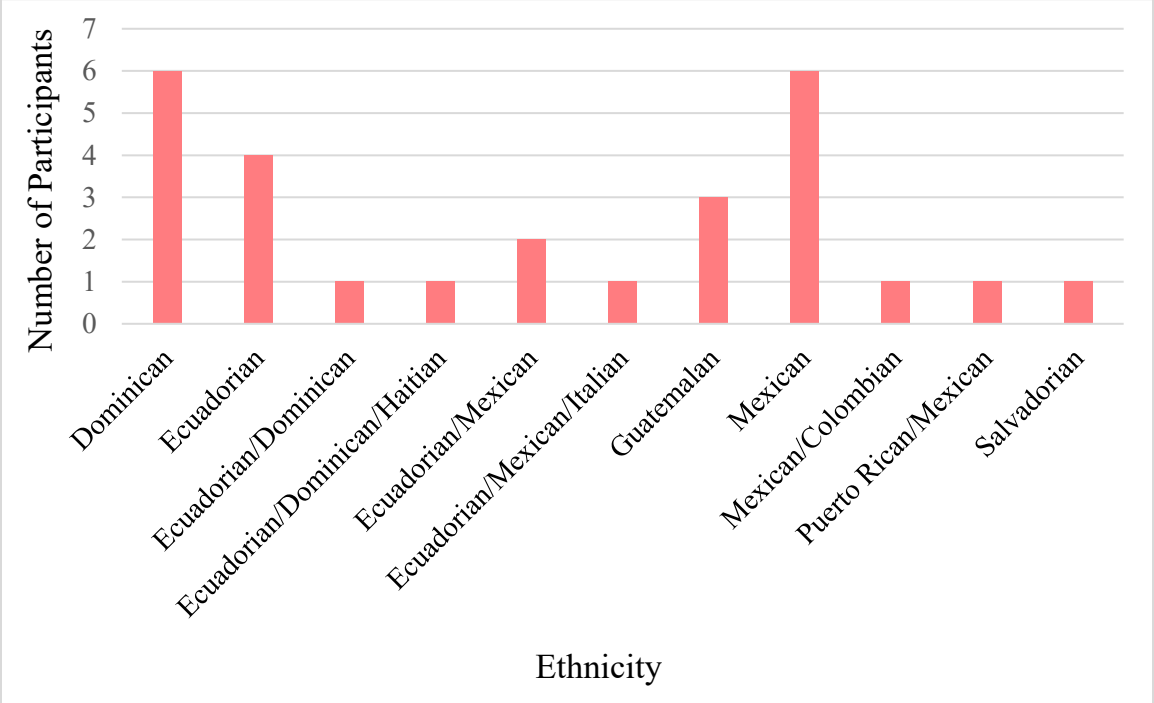
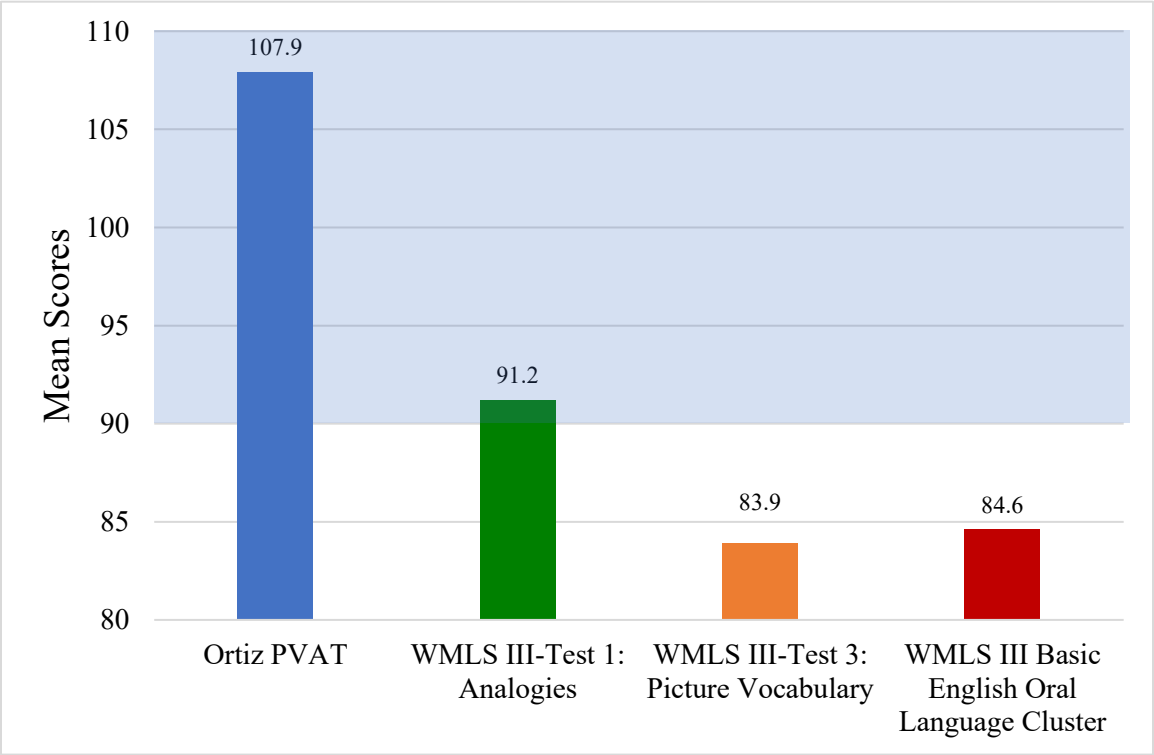


Figure 2

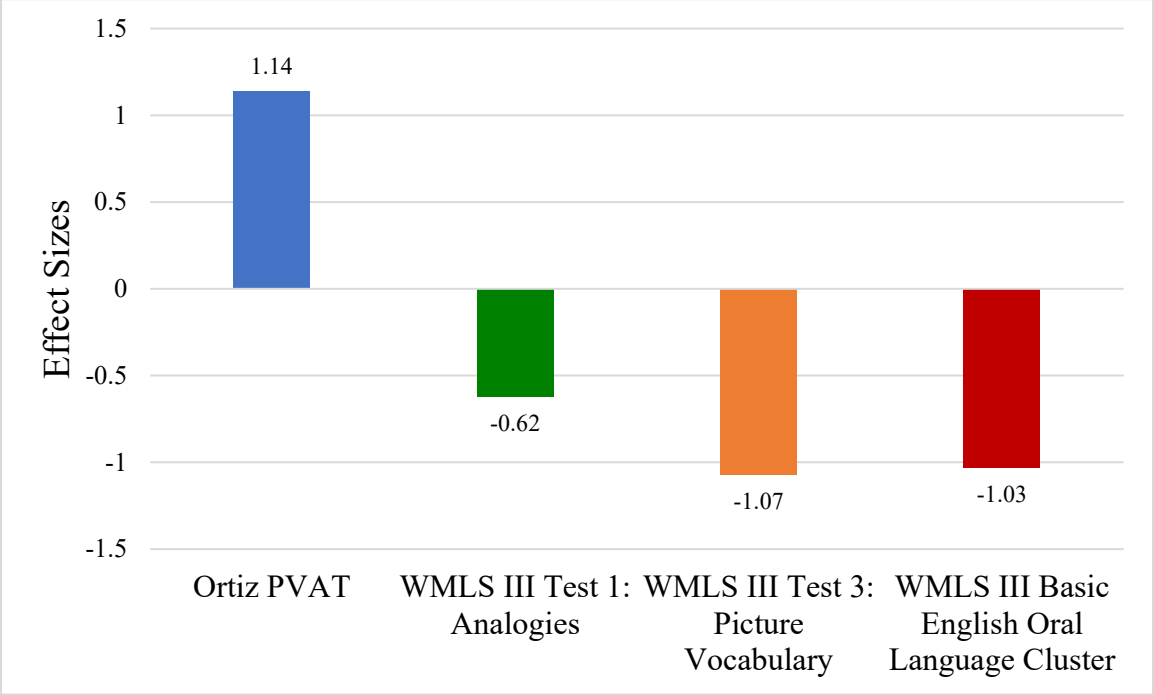
Overall Mean Scores of Ortiz PVAT and WMLS III English



Note. Shaded area indicates standard score average range (90 – 110) for both Ortiz PVAT and WMLS III.

Figure 3

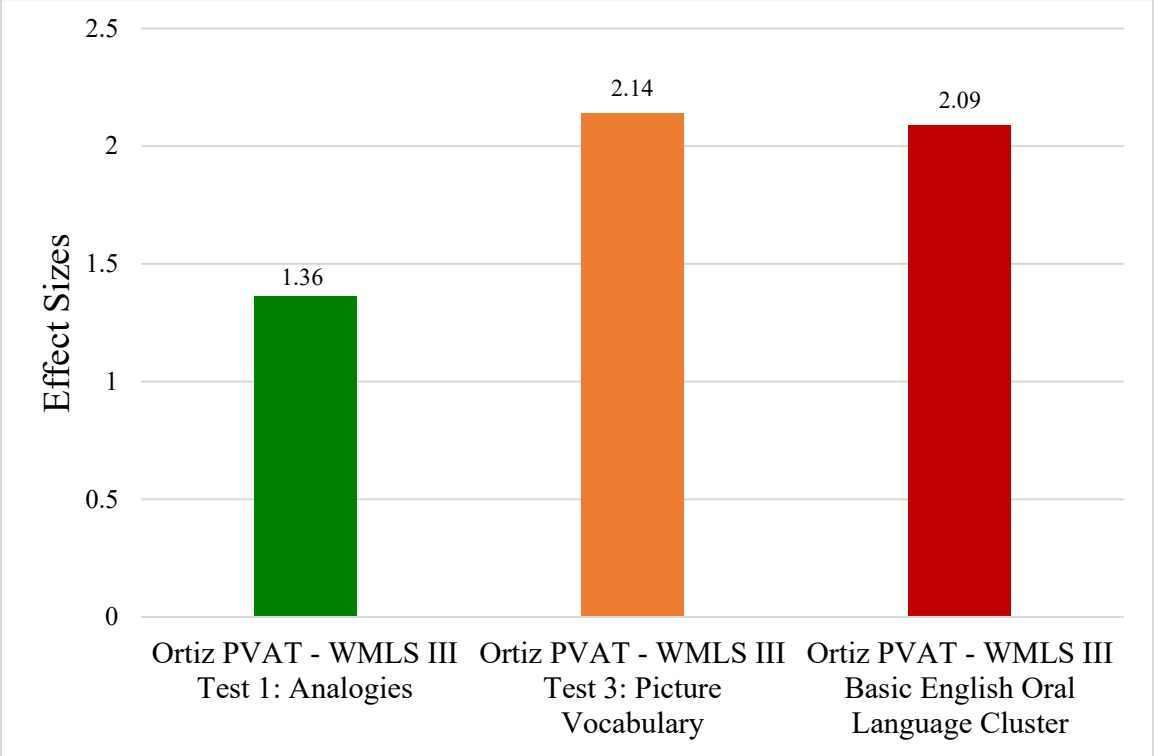
Cohen's d Effect Sizes of Ortiz PVAT and WMLS III English – One Sample



Note. Cohen's *d* effect size = Small .02; Medium = .05; Large = .08; One-Sample T-Tests.

Figure 4

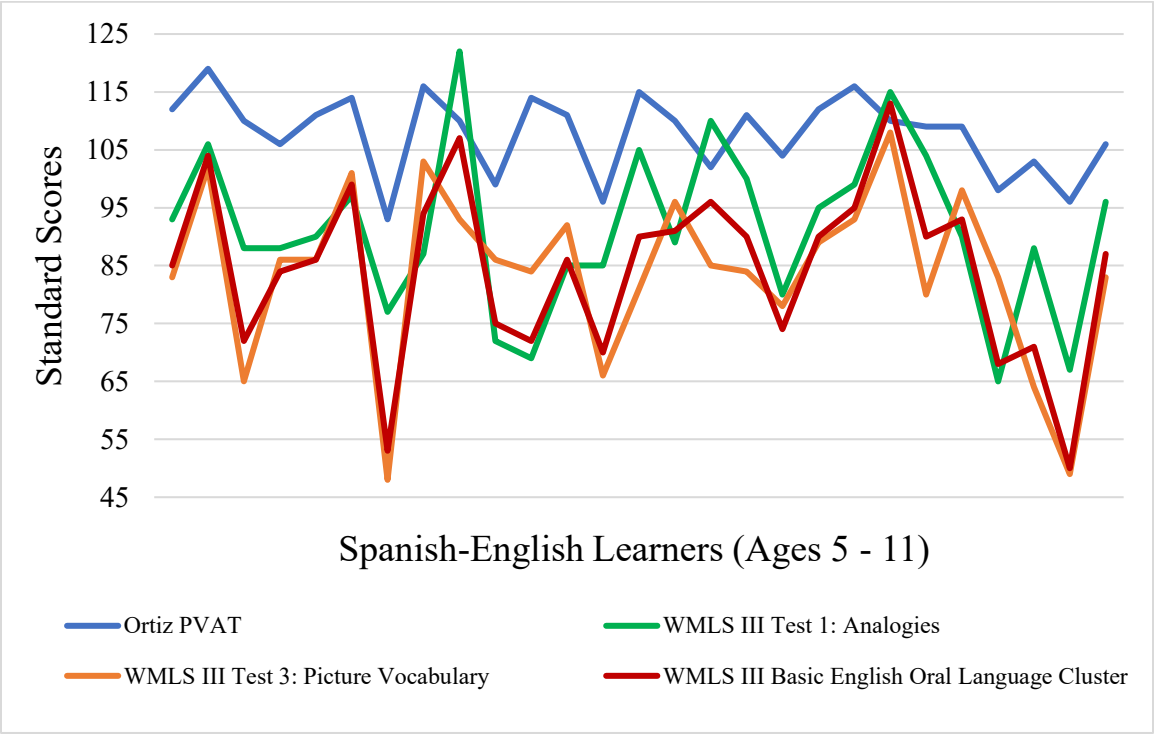
Cohen's d Effect Sizes of Ortiz PVAT and WMLS III English – Paired Samples



Note. Cohen's *d* effect size = Small .02; Medium = .05; Large = .08; Paired Sample T-Tests.

Figure 5

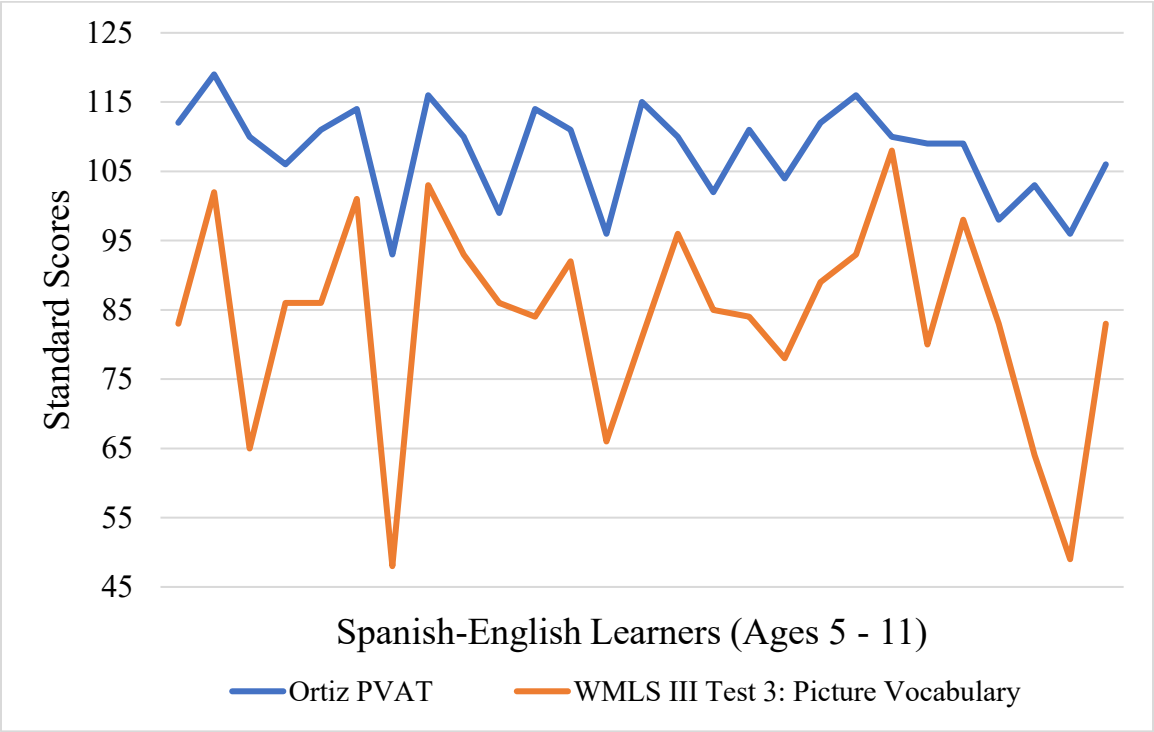
Performance of Spanish-English Learners (Ages 5 – 11) Across the Ortiz PVAT and WMLS III English



Note. One of the participant's data was excluded from the final analysis given the performance scores were an outlier (standard score below 40), which could have affected the analyses.

Figure 6

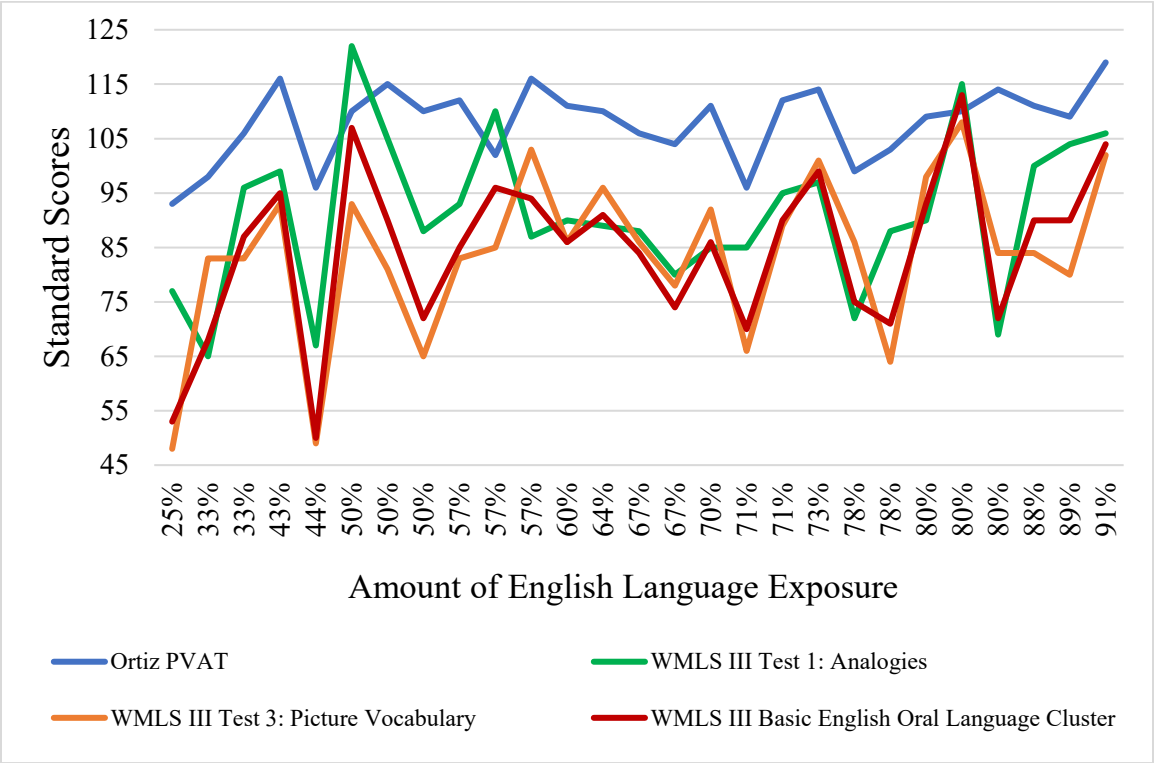
Performance of Spanish-English Learners (Ages 5 – 11) Across the Ortiz PVAT and WMLS III Test 3: Picture Vocabulary



Note. One of the participant's data was excluded from the final analysis given the performance scores were an outlier (standard score below 40), which could have affected the analyses.

Figure 7

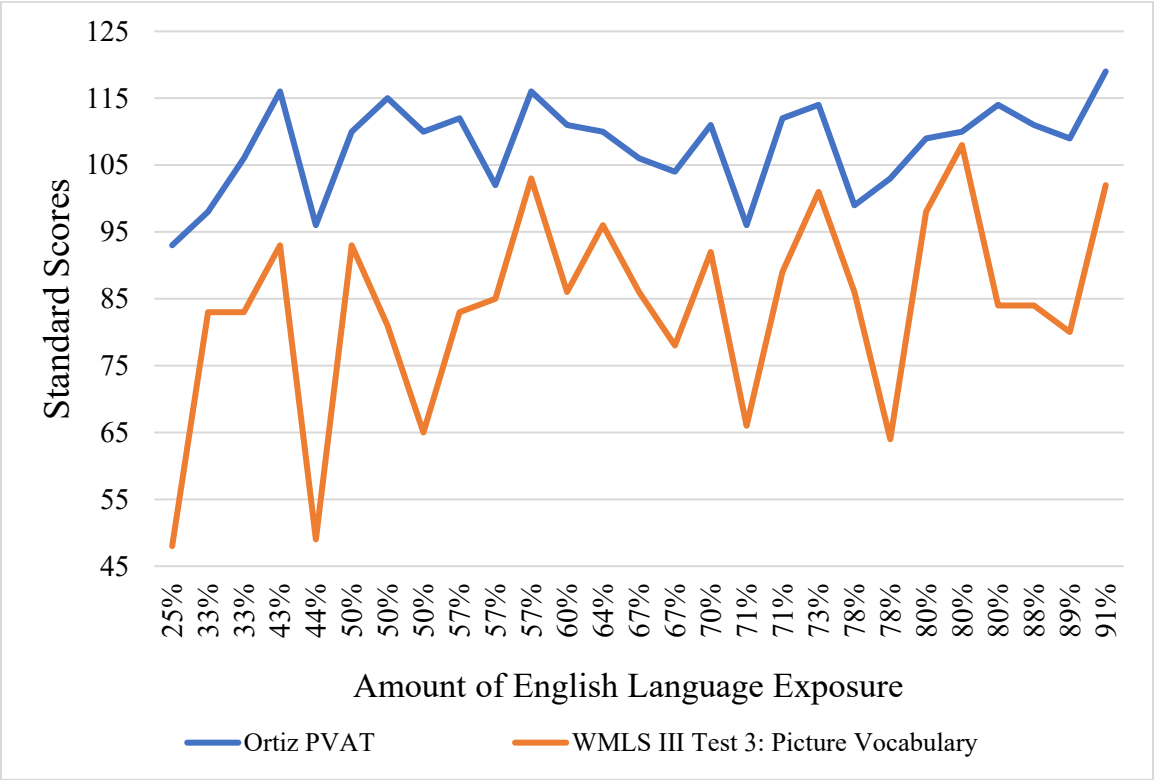
Relationship Between Spanish-English Learners' (Ages 5 – 11) English Language Exposure and Performance in the Ortiz PVAT and WMLS III English



Note. Amount of English exposure is expressed in percent as it relates to each participant's lifetime exposure and was organized from least to greatest. Participants with the same amount of English exposure were placed in order from youngest to oldest.

Figure 8

Relationship Between Spanish-English Learners' (Ages 5 – 11) English Language Exposure and Performance in the Ortiz PVAT and WMLS III Test 3: Picture Vocabulary



Note. Amount of English exposure is expressed in percent as it relates to each participant's lifetime exposure and was organized from least to greatest. Participants with the same amount of English exposure were placed in order from youngest to oldest.

References

Adamek, E. (2019). *Concurrent validity of the WISC-V VCI and the Ortiz PVAT on English learners* (Publication No. 13881450) [Doctoral dissertation, St. John's University]. ProQuest Dissertations Publishing.

American Speech-Language-Hearing Association. (2019). Language in brief. Retrieved from https://www.asha.org/Practice-Portal/Clinical-Topics/Spoken-Language-Disorders/Language-In--Brief/

Arreola v. Santa Ana Board of Education, 476 U.S. 267 (1968).

Atwill, K., Blanchard, J., Christie, J., Gorin, J., & García, H. (2010). English-language learners: Implications of limited vocabulary for cross-language transfer of phonemic awareness with kindergartners. *Journal of Hispanic Higher Education, 9*(2), 104-129.

August, D., Calderón, M., & Carlo, M. (2002). *Transfer of skills from English to Spanish: A study of young leaners – Report for practitioners, parents, and policy makers.* ED-98-CO-0071, Center for Applied Linguistics, Washington: D.C. Retrieved from www.cal.org/acquiringliteracy/pdfs/skills-transfer.pdf

August, D., Carlo, M., Dressler, C., & Snow, C. (2005). The critical role of vocabulary development for English language learners. *Learning Disabilities Research & Practice, 20*(1), 50-57. doi:10.1111/j.1540-5826.2005.00120.x

Berman, P., McLaughlin, B. McLeod, B., Minicuccie, C., Nelson, B., & Woodworth, K. (1995). *School reform and student diversity: Case studies of exemplary practices for LEP students.* Berkeley, CA: National Center for Research on Cultural Diversity and Second Language Learning and BW Associates.

Burger, A., & Chong, I. (2011). Receptive vocabulary. In S. Goldstein & J. A. Naglieri (Eds.), *Encyclopedia of Child Behavior and Development* (pp. 1231-1231). Boston, MA: Springer. DOI https://doi.org/10.1007/978-0-387-79061-9_2359

Byrne, B. M., Oakland, T., Leong, F. T. L., Vijver, F. J. R. van de, Hambleton, R. K., Cheung, F. M., & Bartram, D. (2009). A critical analysis of crosscultural research and testing practices: Implications for improved education and training in psychology. *Training and Education in Professional Psychology, 3*, 94-105. Doi:10.1037/a0014516.

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112* (1), 155-159.

Collier, C. (1989). How long? A synthesis of research on academic achievement in second language. *TESOL Quarterly, 23*, 509-531.

Cummins, J. (2016). Reflections on Cummins (1980), "the cross-lingual dimensions of language proficiency: Implications for bilingual education and the optimal age issue." *Tesol Quarterly, 50*(4), 940-944. doi:10.1002/tesq.339.

Dailor, A. N. (2011). Ethically challenging situations reported by school psychologists: Implications for training. *Psychology in the Schools, 48*(6), 619-631.

Diana v. California State Board of Education, C-70-37 RFP (1970).

Elizalde-Utnick, G., & Romero, P. A. (2017). Assessment of English language learners. In E. C. Lopez, S. G. Nahari, & S. Proctor (Eds.), *Handbook of multicultural school psychology: An interdisciplinary perspective* (pp. 193-217). New York, NY: Routledge.

Endangered Languages Project. (n.d.). *Mam*. Retrieved August 24, 2021, from http://www.endangeredlanguages.com/lang/8091

Endangered Languages Project. (n.d.). *Q'eqchi'*. Retrieved August 24, 2021, from http://www.endangeredlanguages.com/lang/8090

Esquivel, G. B., Warren, T. M., & Littman Olitzky, S. (2007). A historical review of the development of multicultural school psychology. In G. B. Esquivel, E. C. Lopez, & S. G. Nahari (Eds.), *Handbook of multicultural school psychology: An interdisciplinary perspective* (pp. 3-28). Mahwah, NJ: Lawrence Erlbaum Associates.

Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*, 1149-1160.

Figueroa, R. A. (1989). Psychological testing of linguistic-minority students: Knowledge gaps and regulations. *Exceptional Children, 56,* 142-152.

Flanagan, D. P. (2013). *The Cattell-Horn-Carroll theory of cognitive abilities*. Wiley Online Library. Retrieved from https://onlinelibrary.wiley.com/doi/pdf/10.1002/978111866054.e se0431

Flanagan, D. P. & Alfonso, V. C. (2017). *Essentials of WISC-V assessment*. New Jersey: John Wiley & Sons.

Flanagan, D. P., Alfonso, V. C., Ortiz, S. O., & Dynda, A. M. (2010). Best practices in integrating cognitive assessment in school neuropsychological assessments. In D. C. Miller (Eds.), *Best practices in school neuropsychology: Guidelines for effective practice, assessment, and evidence-based intervention* (pp. 101-140). New Jersey: John Wiley & Sons, Inc.

Gámez, P., Griskell, H., Sobrevilla, Y., & Vazquez, M. (2019). Dual language and English-only learners' expressive and receptive language skills and exposure to peers' language. *Child Development, 90*(2), 471-479. doi:10.1111/cdev.13197

Guadalupe v. Tempe Elementary School District, CIV 71-435 U.S. (1972).

Hammer, C. S., Hoff, E., Uchikoshi, Y., Gillandes, C., Castro, D., & Sandilos, L. E. (2014). The language and literacy development of young dual language learners: A critical review. *Early Childhood Research Quarterly, 29*(4), 715-733. doi:10.1016/j.ecresq.2014.05.008

Heinrich-Heine-Universität Düsseldorf. (2020). G*Power: *Statistical power analyses for Windows and Mac*. https://www.psychologie.hhu.de/arbeitsgruppen/allgemeine-psychologie-und-arbeitspsychologie/gpower.html

Houghton Mifflin Harcourt. (2019). Woodcock Muñoz Language Survey – Third Edition (WMLS III). Retrieved from https://www.hmhco.com/programs/woodcock-munoz-language-survey-iii

Huck, S. W. (2012). *Reading Statistics and Research*. 6th Ed. Boston: Allyn and Bacon.

Individuals with Disabilities Education Act (IDEA), 20 U.S.C. Sec. 1400. (2004).

Jones, J. (2014). Best practices in providing culturally responsive interventions. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology VI*. Bethesda, MD: National Association of School Psychologists.

Jongejan, W., Verhoeven, L., & Siegel, L. (2007). Predictors of reading and spelling abilities in first- and second-language learners. *Journal of Educational Psychology, 99*(4), 835-835.

Kangas, S. (2018). Breaking one law to uphold another: How schools provide services to English learners with disabilities. *Tesol Quarterly, 52*(4), 877-910. doi:10.1002/tesq.431

Kiefer, M. J. (2008). Catching up or falling behind? Initial English proficiency, concentrated poverty, and the reading growth of language minority learners in the United States. *Journal of Educational Psychology, 100,* 851-868.

Krashen, S. D. (1982). *Principles and practice in second language acquisition*. Retrieved from http://www.sdkrashen.com/content/books/principles_and_practice.pdf

Lopez, E. C. (2006). English language learners. In G. G. Bear & K. M. Minke (Eds.), *Children Needs III* (pp 647-659). National Association of School Psychologists.

Lopez, E. C., & Bursztyn, A. M. (2013). Future challenges and opportunities: Toward culturally responsive training in school psychology. *Psychology in the schools, 50*(3), 212-228.

Lopez, E. C., & Velasco, P. (2017). Supporting teachers of English learners via instructional consultation. In E. C. Lopez, S. G. Nahari, & S. Proctor (Eds.), *Handbook of multicultural school psychology: An interdisciplinary perspective* (pp. 35-55). New York, NY: Routledge.

Luft Baker, D., Richards-Tutor, C., Gersten, R., Baker, S.K. & Smith, J. L. M. (2017). Building reading skills for English learners within a response to intervention framework. In E. C. Lopez, S. G. Nahari, & S. Proctor (Eds.), *Handbook of multicultural school psychology: An interdisciplinary perspective* (pp. 97-117). New York, NY: Routledge.

Mancilla-Martinez, J., & Lesaux, N. (2011). Early home language use and later vocabulary development. *Journal of Educational Psychology, 103*(3), 535-535.

McLaughlin, B., August, D., Snow, C., Carlo, M., Dressler, C., White, C., Lively, T., & Lippman, D. (2000, April). *Vocabulary improvement and reading in English language learners: An intervention study*. Proceedings of A Research Symposium on High Standards in Reading for Students From Diverse Language Groups: Research, Practice & Policy. Washington, DC: U.S. Department of Education. Retrieved from https://ncela.ed.gov/files/rcd/BE023768/Vocabulary_Improvement.pdf

Nahari, S. G., Martines, D. G., & Wang, P. (2017). Consulting with culturally and linguistically diverse parents. In E. C. Lopez, S. G. Nahari, & S. Proctor (Eds.), *Handbook of multicultural school psychology: An interdisciplinary perspective* (pp. 56-76). New York, NY: Routledge.

National Association of School Psychologists. (2015). Position statement: The provision of school psychological services to bilingual students. Retrieved August 1, 2019, from http://www.nasponline.org/x32086.xml

National Association of School Psychologists. (2017). National Association of School Psychologists Strategic plan: 2017 – 2022. Retrieved August 3, 2019, from https://www.nasponline.org/x41409.xml

National Center on Educational Outcomes. (2016). *ELs with disabilities*. Retrieved from https://nceo.info/student_groups/ells_with_disabilities

National Center for Education Statistics. (2019). *The Condition of Education 2019* (NCES 2019-144). Washington, DC: U.S. Department of Education. Retrieved August 12, 2019, from https://nces.ed.gov/pubs2019/2019144.pdf

Newell, M. L., & Chavez-Korell, S. (2017). The evolution of multiculturalism: An interdisciplinary perspective. In E. C. Lopez, S. G. Nahari, & S. Proctor (Eds.), *Handbook of multicultural school psychology: An interdisciplinary perspective* (pp. 3-17). New York, NY: Routledge.

North Shore Pediatric Therapy. (2021). Expressive vs. receptive language. Retrieved accessed August 24, 2021, from https://www.nspt4kids.com/parenting/expressive-vs-receptive-language/

Ortiz, S. O. (2008). Best practices in nondiscriminatory assessment. In A. Thomas (Eds.) & J. Grimes (Eds.), *Best practices in school psychology V* (pp. 661-678). Bethesda, MD: National Association of School Psychologists.

Ortiz, S. O. (2014). Best practices in nondiscriminatory assessment. In P. L. Harrison & A. Thomas (Eds.), *Best practices in school psychology: Foundations* (pp. 61-74). Bethesda, MD: National Association of School Psychologists.

Ortiz, S. O. (2018a). *Ortiz Picture Vocabulary Acquisition Test*. Multi-Health Systems Inc.

Ortiz, S. O. (2018b). Technical Manual. *Ortiz Picture Vocabulary Acquisition Test*. Multi-Health Systems Inc.

Ortiz, S. O. (2019). On the measurement of cognitive abilities in English learners. *Contemporary School Psychology, 23*(1), 68-86.

Ortiz, S. O., Flanagan, D. P., & Dynda, A. M. (2008). Best practices in working with culturally diverse children and families. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology V* (pp. 1721-1738). Bethesda, MD: National Association of School Psychologists.

Ortiz, S. O., & Wong, J. Y. T. (2020, October). Fairness in tests and test score interpretation with English learners. *The Score*. Retrieved August 24, 2021, from http://www.apadivisions.org/division-5/publications/score/2020/10/assessing-english-learners

Palermo, F., & Mikulski, A. M. (2014). The role of positive peer interactions and English exposure in Spanish speaking preschoolers' English vocabulary and letter-word skills. *Early Childhood Research Quarterly, 29*(4), 625–635. https://doi.org/10.1016/j.ecresq.20 14.07.006

Peña, E. D., Gutiérrez-Clellen, V. F., Iglesias, A., Goldstein, B. A., & Bedore, L. M. (2018). Bilingual English-Spanish Assessment. Baltimore, MD: Brookes Publishing. Retrieved from http://archive.brookespublishing.com/documents/BESA-excerpt.pdf

Rabinowitz, S. (2008). *Assessment of English Language Learners Under Title I and Title III: How One Testing Program Can Inform the Other*. San Francisco, CA: Assessment and Accountability Content Center, LEP Partnership, WestEd.

Rhodes, R. L., Ochoa, S. H., & Ortiz, S. O. (2005). *Assessing culturally and linguistically diverse students: A practical guide.* NY: Guilford.

Schrank, F. A., Miller, D. C., Wendling, B. J., & Woodcock, R. W. (2010). *Essentials of WJ III cognitive abilities assessment*. (2nd ed.). New Jersey: John Wiley & Sons.

Seol, K. I., Song, S. H., Kim, K. L., Oh, S. T., Kim, Y. T., Im, W. Y., Song, D. H., & Cheon, K. A. (2014). A comparison of receptive-expressive language profiles between toddlers with autism spectrum disorder and developmental language delay. *Yonsei Medical Journal*, *55*(6), 1721–1728. https://doi.org/10.3349/ymj.2014.55.6.1721

Skiba, R. J., Simmons, A. B., Ritter, S., Gibb, A. C., Rausch, M. K., Cuadrado, J., & Chung, C. G. (2008). Achieving equity in special education: History, status, and current challenges. *Council for Exceptional Children, 74*(3)*,* 264-288.

Tello, A. C. (2020). *Comparison of language performance using exposure and non-exposure norms in English learners identified with speech language impairment* (Publication No. 28086833) [Doctoral dissertation, St. John's University]. ProQuest Dissertations Publishing.

Uchikoshi, Y. (2006). English vocabulary development in bilingual kindergarteners: What are the best predictors? *Bilingualism, 9*(1), 33-49.

United States Census Bureau. (2018). More than 76 million students enrolled in U.S. schools, Census bureau reports. Retrieved August 1, 2019, from https://www.census.gov/newsroom/press-releases/2018/school-enrollment.html

U.S. Department of Education (n.d.). Our nation's English learners. Retrieved November 1, 2019, from https://www2.ed.gov/datastory/el-characteristics/index.html#two

U.S. Department of Education. (2015). Ensuring English learner students can participate meaningfully and equally in educational programs. Retrieved August 12, 2019, from https://www2.ed.gov/about/offices/list/ocr/docs/dcl-factsheet-el-students-201501.pdf

Urbaniak, G.C., & Plous, S. (2021). Research randomizer (Version 4.0) [Computer software]. Retrieved from https://www.randomizer.org/

Verdugo, R., & Flores, B. (2007). English-language learners: Key issues. *Education and Urban Society, 39*(2), 167-193. doi:10.1177/0013124506294852

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

Woodcock, R. W., Alvarado, C. G., & Ruef, M. L. (2017). *Woodcock-Muñoz Language Survey III.* Itasca, IL: Houghton, Mifflin Harcourt.

Woodcock, R. W., Alvarado, C. G., Ruef, M. L., & Schrank, F. A. (2017). Comprehensive Manual. *Woodcock-Muñoz Language Survey III.* Itasca, IL: Houghton, Mifflin Harcourt.

Yzquierdo, Z., Blalock, G., & Torres-Velasquez, D. (2004). Language-appropriate assessments for determining eligibility of English language learners for special education services. *Assessment for Effective Intervention, 29*(2), 17-30.

Zoom. (2021). Zoom Video Communications. (Version 5.7.4) [Computer software]. Retrieved from https://zoom.us/download

Vita

Name	*Amanda García*
Baccalaureate Degree	*Bachelor of Arts, Queens College (CUNY), Flushing, New York* *Major: Psychology*
Date Graduated	*May, 2010*
Other Degrees and Certificates	*Master of Science, Queens College (CUNY), Flushing, New York* *Major: School Psychology* *Bilingual Extension, Queens College (CUNY)*
Date Graduated	*May, 2015*